ZEUS PRESS

Financial Economics Research

VOLUME 3 · ISSUE 4 · 2026

E-ISSN 3006-3663 · P-ISSN 3006-3655

© 2026 The Author(s). Authors retain copyright in their articles.

ISSN 3006-3655
9 773006 365008 04>

FINANCIAL ECONOMICS RESEARCH · JOURNAL · VOLUME 3 · ISSUE 4 · 2026

ZEUS PRESS

FER

Financial Economics Research

OPEN ACCESS · DOUBLE-BLIND PEER-REVIEWED ACADEMIC JOURNAL

VOLUME 3 · ISSUE 4 · 2026

E-ISSN 3006-3663 · P-ISSN 3006-3655

https://journals.zeuspress.org/index.php/fer/index

ZEUS PRESS

Financial Economics Research

FER

ISSUE
Volume 3 · Issue 4 · 2026

PUBLISHED BY
Zeus Press

OPEN ACCESS · DOUBLE-BLIND PEER-REVIEWED ACADEMIC JOURNAL

VOLUME 3 · ISSUE 4 · 2026

E-ISSN 3006-3663 · P-ISSN 3006-3655 · DOI Prefix 10.70267

https://journals.zeuspress.org/index.php/fer/index

Recommended citation: Financial Economics Research, Volume 3, Issue 4 (2026).

Financial Economics Research

ISSN	3006-3663 (Online), 3006-3655 (Print)
DOI PREFIX	10.70267
WEBSITE	https://journals.zeuspress.org/index.php/fer/index
CONTACT	service@zeus-press.org; +852 56089963

ABOUT THE JOURNAL

Financial Economics Research (FER) is a double-blind peer-reviewed, open-access journal that publishes high-quality research in the field of financial economics. The journal aims to provide a platform for scholars, researchers, and practitioners to share cutting-edge research and insights that contribute to the understanding and enhancement of financial markets, institutions, and practices.

TYPES OF SUBMISSION

The journal publishes: (1) full-length articles, (2) review articles, (3) book reviews, (4) discussion notes and (5) articles in translation.

EDITORIAL AND PEER-REVIEW POLICY

All research articles are subject to editorial assessment and double-blind peer review. The journal follows recognized standards of publication ethics, research integrity, transparency and correction.

PUBLICATION ETHICS AND INTEGRITY

The journal applies principles of research integrity, transparent authorship, disclosure of competing interests, data accuracy and prompt correction of the scholarly record. Suspected misconduct is assessed through an impartial editorial process.

AUTHOR RESPONSIBILITIES

Authors retain responsibility for the accuracy, originality and ethical compliance of their work, including acknowledgements, funding statements, data availability and conflict-of-interest disclosures.

PERMISSIONS

The author is responsible for obtaining all permissions required before submission. Permission and owner details must be provided for third-party content, methods, software, questionnaires and scales. Authors must check applicable licences, obtain permission where required, and include confirmation in the Materials and Methods section.

OPEN ACCESS LICENCE

This publication is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits noncommercial use, sharing, adaptation, distribution and reproduction with appropriate credit, a link to the licence, and an indication of changes.

DIGITAL PRESERVATION AND RECORD

The publisher supports long-term accessibility of the scholarly record through persistent identifiers, structured metadata and established digital preservation practices. Version and correction information should remain linked to the article of record.

© 2026 The Author(s). Authors retain copyright in their articles.

Published by: Zeus Press

Address: Flat/Rm 91_3 9/F Chinachem Golden Plaza No.77 Mody Road Tsim Sha Tsui KL, Hong Kong

Individual paper reprints: service@zeus-press.org

CONTENTS

Financial Economics Research · Volume 3 · Issue 4 · 2026

Green Finance Policy and Corporate Green Innovation: Empirical Evidence from Green Finance Reform and Innovation Pilot Zones 1
Zhengyu Luo, Wenjing Yang

Green Finance and Corporate Green Innovation 18
Xinyu Wen

Research on the Impact of High-Speed Rail Opening on Financial Resource Allocation in the Context of Urban-Rural Integration: An Estimation Based on the PSM-DID Model 33
Jiaying Li

Analysis of the Impact of Financial Technology on the Liquidity Risk of Commercial Banks 47
Mingrui Su

Tax Incentives and Debt-Financing Efficiency in Advanced Manufacturing Enterprises: Evidence from Chinese A-Share Listed Firms 53
Pengfei Qian

Executive Overseas Background, Customer Concentration, and Firm Internationalization: Evidence from Structural Changes Around 2018 61
Zikun Wu

The Spatial and Temporal Distribution Characteristics and Governance Tension of Tricycle Passenger Transport in County Towns from the Perspective of Informal Economy—A Case Study of A Certain County in Guangdong Province 67
Yue Li

Can Annual Report Tone Signal Audit Risk? Evidence from Large Language Model Textual Analysis of Chinese Listed Firms 74
Wen Guan, Yuxuan Zhou, Jing Zhang, Junyi Zang

Cross-Sectional Return Prediction in China's A-Share Market Based on Lasso and XGBoost: A Comparison with the Fama–MacBeth Multi-Factor Model 84
Xichun Zhang

Science and Technology Finance Pilot Policy and Corporate Long-term Investment: Evidence from China 97
Yunqi Lyu, Min Gao, Kunyang Xue

Geopolitical Shift in Capital Preference: The Differential Impact of Geopolitical Risk and Political Fragmentation on Cross-Border Capital Allocation 111
Yong Zhou

VAT Credit Refunds and Firms' Export Development Evidence from Chinese A-Share Listed Companies 118
Ken Peng

CONTENTS

Financial Economics Research · Volume 3 · Issue 4 · 2026

The Study of the Impact Mechanism of R&D Expense Super-deduction on Green Innovation of High-tech Enterprises in the Context of the Digital Economy 129
Ting Wang

The Evolution of Predictive Analytics: Machine Learning and the Predictive Step in Empirical Asset Pricing—From Factor Zoo to Factor Refinery 143
Kaiyu Wu

Research on the Mechanism in the Transformation of BYD's Strategic Advantages into Financial Performance: Based on the Harvard Analysis Framework 162
Yuyue Zhang

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.0117
Published by: Zeus Press

Green Finance Policy and Corporate Green Innovation: Empirical Evidence from Green Finance Reform and Innovation Pilot Zones

Zhengyu Luo[1,*,†] and Wenjing Yang[2,†]

[1]*School of Economics, Shandong University, Jinan 250100, China*

[2]*Bay Area International Business School, Beijing Normal University, Zhuhai 519087, China*

**Corresponding author: Zhengyu Luo.*

† *These authors contributed equally to this work.*

Abstract

Green innovation is characterized by both dual externalities and long cycles, and generally faces financing constraints and governance deficiencies. Relying solely on market mechanisms is insufficient to fully unleash the vitality of corporate green innovation. Using Shanghai and Shenzhen A-share listed companies from 2014 to 2024 as the research sample and leveraging the staggered establishment of green finance reform and innovation pilot zones as a quasi-natural experiment, this paper constructs a multi-period difference-in-differences (DID) model to systematically examine the micro-level incentive effects and transmission mechanisms of the pilot zone policies, analyzing them through the dual dimensions of innovation intensity and innovation willingness. Empirical results indicate that the policy establishing pilot zones for green finance reform and innovation significantly boosts the level of green innovation among enterprises within these jurisdictions. Specifically, the volume of green patent applications by pilot enterprises increased by an average of 12.92% compared to non-pilot enterprises. The policy exhibits both a corporate governance substitution effect and an information disclosure extension effect, demonstrating a more pronounced promotional impact on innovation among enterprises with weak internal governance foundations or inadequate environmental information disclosure, thereby reflecting a mechanism that addresses and remedies specific weaknesses. The marginal contributions of this paper are threefold: It first replaces the focus on single green financial instruments with comprehensive policy pilots, employs a multi-period difference-in-differences (DID) approach to mitigate endogeneity issues, and refines the measurement system for green innovation across dual dimensions. It also transcends the limitations of existing literature, which largely focuses on the financing constraint channel, by revealing the dual functions of green finance in resource allocation and external governance. Finally, it characterizes the heterogeneous effects of the policy across different types of firms, providing micro-level empirical evidence to support targeted policy implementation and differentiated empowerment of corporate green innovation.

Keywords

green finance reform and innovation pilot zones, green finance, corporate green innovation, multi-period difference-in-differences, corporate governance, information disclosure

© 2026 The Author(s). Authors retain copyright in their articles.

1. Introduction

Against the backdrop of cultivating new quality productive forces and advancing a comprehensive green transformation of the economy and society, green technology innovation serves as the core driving force for achieving "dual carbon" goals, overcoming resource and environmental constraints, and upgrading industrial structures. However, green innovation is characterized by dual externalities, namely, technology spillovers and positive environmental externalities, meaning that, under market equilibrium, corporate investment in innovation often falls short of the socially optimal level [1]. At the same time, green R&D is characterized by large-scale investment, long payback periods, and high technical uncertainty. Furthermore, its outputs are primarily intangible assets that lack the physical collateral required for traditional credit, causing enterprises to generally face significant financing and risk constraints [2,3]. It is difficult to fully mobilize the vitality of green innovation by relying solely on spontaneous market regulation. Institutional policy guidance and financial system support are needed.

Green finance is a key lever for supply-side structural reform in the financial sector and a core institutional arrangement for driving green development. Following the 2016 release of the Guidelines on Establishing the Green Financial System, China's green finance entered a stage of systematic advancement driven by top-level design. To explore local practice models that can be replicated and scaled up, the State Council established green finance reform and innovation pilot zones in three batches starting in 2017, creating a policy framework that employs a dual approach of incentives and constraints. It has provided an ideal quasi-natural experiment setting for evaluating the microeconomic effects of green finance.

Existing literature has extensively explored the microeconomic effects of green finance. Regarding policy impact, most studies focus on single policy instruments, such as the Green Credit Guidelines, and find that they exert a significant "forcing effect" on green innovation among heavily polluting enterprises [4, 5]. Some studies have also pointed out that tightening financing constraints may squeeze out corporate R&D resources [6,7]. But there is no consensus on the conclusion. Regarding comprehensive policy pilot programs such as green finance reform and innovation pilot zones, existing studies have largely focused on their impacts on corporate financing costs, pollution emissions, and investment structures [8,9]. However, it remains difficult to accurately capture the marginal impact of these policies on corporate innovation behavior. Regarding transmission mechanisms, most scholars consider financing constraints to play a mediating role [10, 11], yet the governance attributes of green finance policies are often overlooked. Another potential pathway involves mitigating information asymmetry and enhancing the precision of green capital allocation through strengthened information disclosure systems. Yet neither of these transmission pathways has been fully tested or validated.

This paper selects Shanghai and Shenzhen A-share listed companies from 2014 to 2024 as the research sample. Leveraging the staggered establishment of pilot zones as a quasi-natural experiment, it constructs a multi-period difference-in-differences (DID) model to systematically examine the incentive effects of green finance policies on the intensity and willingness of corporate green innovation. Furthermore, it conducts cross-sectional analyses across the dimensions of corporate governance and information disclosure to reveal the policy transmission pathways and stratified characteristics.

This paper makes three primary contributions. By focusing on comprehensive policy pilots and employing a staggered difference-in-differences (DID) approach to mitigate endogeneity, it refines the measurement system for green innovation, incorporating both innovation intensity and willingness, to more accurately identify the policy's causal effects. Moreover, building on existing literature that emphasizes financing constraints, it verifies the dual functions of green finance, namely, resource allocation and external governance, thereby enriching research into the transmission mechanisms through which green finance fosters innovation. Finally, it characterizes the heterogeneity of policy effects across dimensions such as internal governance foundations, environmental management systems, and information disclosure levels. In doing so, it reveals the non-linear nature of the policy's role in "addressing shortcomings," providing micro-level empirical evidence to support the targeted empowerment of corporate green innovation.

The remainder of this paper is organized as follows: Section 2 outlines the institutional background and research hypotheses, reviewing the institutional evolution of the green finance reform and innovation pilot zones and formulating hypotheses based on theoretical analysis. Section 3 details the research design, including sample selection, data sources, model specification, and variable definitions. Section 4 presents the empirical

© 2026 The Author(s). Authors retain copyright in their articles.

results and analysis, reporting baseline regression findings, validity tests for the Difference-in-Differences (DID) approach, and a series of robustness checks. Section 5 conducts an analysis of cross-sectional heterogeneity, examining policy transmission channels through the lenses of corporate governance and information disclosure. Section 6 offers research conclusions and policy recommendations.

2. Institutional Background and Research Hypotheses

2.1 Institutional Evolution of Pilot Zones for Green Finance Reform and Innovation

In August 2016, seven government agencies, including the People's Bank of China, jointly issued the Guiding Opinions on Establishing a Green Financial System. This marked the first time the concept of green finance was defined, an institutional framework established, and core tasks clarified at the national level, signaling the entry of China's green finance development into an institutionalized phase guided by top-level design.

To overcome bottlenecks in the local implementation of green finance, China adopted a reform strategy characterized by phased piloting and a tiered rollout. In June 2017, the State Council approved the establishment of the first batch of pilot zones in Huzhou and Quzhou (Zhejiang), Ganjiang New Area (Jiangxi), Guangzhou (Guangdong), Gui'an New Area (Guizhou), and Hami, Changji, and Karamay (Xinjiang), with each location exploring approaches tailored to its specific industrial characteristics and resource endowments. In November 2019, Lanzhou New Area (Gansu) was approved as a second-batch pilot zone, focusing on ecological protection in the Yellow River Basin and new energy industry clusters. In August 2022, the entirety of Chongqing was designated as a third-batch pilot zone, which marks the first province-level pilot covering an entire administrative region, thereby advancing the reform from isolated cities toward regional coordination.

2.2 Theoretical Analysis and Research Hypotheses

Green innovation is characterized by long cycles, high uncertainty, and pronounced dual externalities. Corporate investment in innovation is constrained by both the external financing environment and the quality of internal corporate governance. Green projects involve complex technical pathways and long timelines for realizing returns. Significant information friction exists between enterprises and financial institutions, and green R&D has long faced a shortage of financing [12]. Green innovation entails long-term strategic investment. However, pressure from short-term performance evaluations and agency conflicts can foster short-sighted behavior, such as the voluntary reduction of green R&D budgets, while internal governance deficiencies further diminish the efficiency of innovation resource allocation [13].

Green finance reform and innovation pilot zones serve as comprehensive policy instruments capable of empowering corporate green innovation through two pathways: first, by broadening low-cost financing channels, such as green credit and green bonds, to alleviate funding gaps; and second, by leveraging environmental credit assessments and mandatory information disclosure to establish external oversight, thereby compelling enterprises to increase investment in green R&D. Based on this, the following baseline hypothesis is proposed:

H1: Holding other factors constant, the policies for green finance reform and innovation pilot zones can significantly enhance the intensity of and willingness for green innovation among enterprises within the jurisdiction.

From the perspective of internal governance, enterprises with robust internal controls and ISO 14001 certification have established mature green operational mechanisms and possess strong momentum for independent innovation. In contrast, enterprises with weak governance suffer from issues such as the misuse of funds and a lack of environmental management. The pilot zones link credit resources with environmental compliance, use external constraints to make up for internal governance shortcomings, and have a stronger incentive effect on companies that address shortcomings. It is therefore proposed that:

H2: Policies for green finance reform and innovation pilot zones exhibit a corporate governance substitution effect, exerting a stronger promoting influence on green innovation in enterprises with lower levels of internal governance.

© 2026 The Author(s). Authors retain copyright in their articles.

From the perspective of information disclosure, companies with comprehensive disclosure regarding environmental principles and social responsibility exhibit lower levels of information asymmetry, making it easier for their green projects to secure funding. It is difficult for financial institutions to recognize the green value of enterprises with weak information disclosure. By standardizing disclosure norms and establishing a green credit system, the pilot zones have significantly improved the financing environment for enterprises with low disclosure levels, yielding a stronger marginal effect in boosting innovation; thus, the following hypothesis is proposed:

H3: The green finance reform and innovation pilot zone policy has an information disclosure extension effect and has a stronger promotion effect on green innovation of enterprises with a weak information disclosure foundation.

3. Research Design

3.1 SAMPLE Selection and Data Sources

This study focuses on Shanghai and Shenzhen A-share listed companies from 2014 to 2024. Samples were selected based on the following criteria: excluding companies in the finance, insurance, and real estate sectors; excluding those subject to special treatment (ST/SZ) or abnormal trading status; and removing observations with abnormal operating conditions or missing key data. To mitigate the impact of extreme values on empirical results, all continuous variables were winsorized at the 1% level (two-tailed). Firm- and city-level data were obtained from the CSMAR database and the CNRDS platform, resulting in a balanced panel dataset for subsequent analysis.

3.2 Model Specification and Variable Definitions

To identify the causal effect of green finance policies on corporate green innovation, this paper constructs a multi-period difference-in-differences (Multi-period DID) model, leveraging the quasi-natural experiment created by the phased establishment of green finance reform and innovation pilot zones. The baseline model is specified as follows:

$$Inno_{i,t}=\beta_0+\beta_1 DID_{i,t}+\gamma X_{i,t}+\delta_i+\lambda_t+\varepsilon_{i,t} \quad (1)$$

The core explanatory variable, the treatment effect variable ($DID_{i,t}$) associated with the implementation of the green finance reform and innovation pilot zone policy, is used to reflect the net impact of green finance policies. The assignment rule for this variable is as follows: it is assigned a value of 1 if the enterprise is registered in a green finance reform and innovation pilot zone and the observation year is subsequent to the year the policy was approved; otherwise, it is assigned a value of 0. In the model, $X_{i,t}$ represents firm-level control variables used to account for the effects of individual firm characteristics, such as firm size, leverage ratio, profitability, growth potential, and ownership structure. δ_i is used as a firm-level fixed effect to capture the inherent characteristics of firms that remain constant over time across the various stages. λ_t represents an individual-specific fixed random effect that captures common time trends resulting from nationwide macroeconomic shocks and cyclical fluctuations across industries. $\varepsilon_{i,t}$ is the firm-specific random disturbance term. The coefficient β_1 represents the average treatment effect of the green finance reform and innovation pilot zone policy on corporate green innovation. If β_1 is significantly positive, it indicates that the green finance policy exerts a significant promoting effect on green innovation among enterprises within the pilot zones.

The dependent variable, the intensity of corporate green innovation, is used to measure the strength of green innovation. The enterprise's willingness to engage in green innovation is a binary dummy variable used to measure whether the enterprise conducts green innovation activities. This study also selects other key factors potentially influencing the intensity of corporate green innovation as control variables; these include firm-level variables such as asset size, asset-liability ratio, growth capacity, return on net assets, cash flow ratio, years since listing, and the shareholding ratio of the largest shareholder.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 1: Variable Names and Definitions

Variable Type	Variable Name	Symbol	Measurement / Definition
Dependent Variables	Corporate Green Innovation Intensity	*Inno*	ln (number of green patent applications by the firm in the year + 1)
	Corporate Green Innovation Willingness	*PatentH*	1 if Inno ≠ 0; 0 if Inno = 0
Explanatory Variable	Treatment effect of the Green Finance Reform and Innovation Pilot Zone policy	*DID*	1 if the firm's registered location is in a Green Finance Reform and Innovation Pilot Zone and the observation year is after the policy approval year; 0 otherwise
Control Variables	Firm Size	*Size*	ln (year-end total assets)
	Leverage	*Lev*	Year-end total liabilities / year-end total assets
	Return on Equity	*Roe*	Net profit / year-end net assets
	Cash Flow	*Cfo*	Net cash flow from operating activities / total assets
	Operating Revenue Growth Rate	*Growth*	Current year's operating revenue / previous year's operating revenue − 1
	Years Listed	*Age*	Current year − listing year
	Board Size	*Board*	ln (number of board members)
	Independent Director Ratio	*Indep*	Number of independent directors / total number of directors
	Ownership Nature	*SOE*	1 if the firm is a state-owned enterprise; 0 otherwise
	Largest Shareholder Ownership	*Top1*	Shares held by the largest shareholder / total shares

4. Empirical Results and Analysis

4.1 Descriptive Statistical Results and Analysis

Table 2 shows the descriptive statistical results of the main variables in this article. In terms of explained variables, the mean value of corporate green innovation intensity is 0.985, and the standard deviation is 1.231, indicating that there are certain differences in the level of green patent applications among sample companies The average value of corporate green innovation willingness is 0.520, indicating that approximately 52.0% of the sample companies have carried out green patent application activities. In terms of explanatory variables, the mean value of the policy interaction term (DID) in the Green Finance Reform and Innovation Pilot Zone is 0.047, and the standard deviation is 0.211, which is basically consistent with the statistical characteristics of related policy impact studies.

Among the control variables, the average company size is 22.339, the average asset-liability ratio is 0.424, the average return on net assets is 0.040, the average operating cash flow is 0.048, and the average operating income growth rate is 0.133. The average listing years is 11.729 years, the average board size is 2.102, the average proportion of independent directors is 37.885%, the average property rights nature is 0.320, and the average shareholding ratio of the largest shareholder is 34.622%. The distribution characteristics of the above variables are basically consistent with the statistical results of Chinese listed companies in the existing literature, indicating that the sample in this article is well representative.

Table 2: Descriptive Statistics

	N	Mean	SD	Min	p25	Median	p75	Max
Inno	37010	0.985	1.231	0.000	0.000	0.693	1.792	4.787
PatentH	37010	0.520	0.500	0.000	0.000	1.000	1.000	1.000
DID	37010	0.047	0.211	0.000	0.000	0.000	0.000	1.000
Size	37010	22.339	1.287	19.964	21.419	22.145	23.067	26.206
Lev	37010	0.424	0.207	0.060	0.260	0.413	0.571	0.947
Roe	37010	0.040	0.167	-0.927	0.020	0.062	0.108	0.411
Cfo	37010	0.048	0.067	-0.156	0.010	0.047	0.087	0.242
Growth	37010	0.133	0.389	-0.614	-0.054	0.078	0.232	2.335
Age	37010	11.729	8.181	0.942	4.597	9.722	18.497	29.569
Board	37010	2.102	0.197	1.609	1.946	2.197	2.197	2.639
Indep	37010	37.885	5.371	33.330	33.330	36.360	42.86	57.14
SOE	37010	0.320	0.466	0.000	0.000	0.000	1.000	1.000
Top1	37010	34.622	15.022	7.020	23.220	32.690	44.929	74.34

© 2026 The Author(s). Authors retain copyright in their articles.

4.2 Univariate Statistical Results and Analysis

Table 3 shows the T-test results of the mean differences of each variable before and after the implementation of the Green Finance Reform and Innovation Pilot Zone policy. It can be seen from the figure that after the implementation of the policy, the average value of corporate green innovation intensity and corporate green innovation willingness are higher than before the policy implementation, and they are both significant at the 1% level, indicating that the level of corporate green innovation has been improved after the implementation of the policy. In addition, before and after the implementation of the policy, the asset-liability ratio, return on net assets, operating cash flow, operating income growth rate, board size, proportion of independent directors, nature of property rights and shareholding ratio of the largest shareholder are also different, which means that certain attributes of the company have changed significantly before and after the implementation of the policy. Univariate test results show that there are significant differences in the asset-liability ratio, profitability level, operating cash flow, ownership structure and other characteristics of enterprises before and after the implementation of the policy, indicating that the above time-varying factors may be related to policy impact and corporate innovation behavior. Therefore, subsequent benchmark regressions will include these time-varying control variables to alleviate omitted variable bias. This article will control year fixed effects, industry fixed effects and company fixed effects in regression analysis to reduce the interference of other omitted variables.

Table 3: Univariate Statistics

Variable	Before Policy		After Policy		Mean Difference
	Observations	Mean	Observations	Mean	
Inno	35284	0.980	1726	1.081	-0.101***
PatentH	35284	0.517	1726	0.574	-0.056***
Size	35284	22.337	1726	22.382	-0.045
Lev	35284	0.424	1726	0.416	0.009*
Roe	35284	0.040	1726	0.048	-0.007*
Cfo	35284	0.048	1726	0.057	-0.009***
Growth	35284	0.135	1726	0.105	0.030***
Age	35284	11.740	1726	11.514	0.226
Board	35284	2.103	1726	2.089	0.013***
Indep	35284	37.865	1726	38.298	-0.432***
SOE	35284	0.318	1726	0.362	-0.044***
Top1	35284	34.532	1726	36.457	-1.926***

4.3 Benchmark Regression Results and Analysis

Table 4 shows the baseline regression results of this article. The explained variables in columns (1) and (2) are corporate green innovation intensity, and the explained variables in columns (3) and (4) are corporate green innovation intentions.

In column (1), when only controlling for firm and year fixed effects, the coefficient of the green finance reform and innovation pilot zone policy interaction term DID is 0.1409, which is significantly positive at the 5% level. In Column (2), after including all control variables, the DID coefficient remains 0.1292 and is significantly positive at the 5% level, indicating that the implementation of the green finance reform and innovation pilot zone policy has significantly boosted corporate green innovation efforts.

In column (3), when only company and year fixed effects are controlled, the coefficient of DID is 0.0489, which is significantly positive at the 5% level. After adding control variables to column (4), the coefficient of DID is 0.0461, which is significantly positive at the 10% level, indicating that policy implementation also significantly increases the willingness of enterprises to carry out green innovation activities.

As can be seen from column (2) in the table, the coefficient of the policy interaction term DID is 0.1292, indicating that compared with similar companies in non-pilot areas, the green patent application level of enterprises in pilot areas has increased by approximately 12.92% on average, which has relatively strong economic implications. The above results prove that the research hypothesis of this article is correct, that is, the policies of the green financial reform and innovation pilot zone can improve the green innovation of enterprises.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 4: Benchmark Regression Results

	Inno		PatentH	
	(1)	(2)	(3)	(4)
DID	0.1409**	0.1292**	0.0489**	0.0461*
	(2.34)	(2.23)	(1.96)	(1.85)
Size		0.3802***		0.1180***
		(19.28)		(16.72)
Lev		-0.0799		-0.0161
		(-1.47)		(-0.63)
Roe		-0.0004		-0.0158
		(-0.01)		(-1.06)
Cfo		-0.0175		0.0166
		(-0.24)		(0.43)
Growth		0.0019		0.0086
		(0.18)		(1.54)
Age		-0.0049		-0.0042
		(-0.18)		(-0.28)
Board		0.1144*		0.0560**
		(1.74)		(1.98)
Indep		0.0049**		0.0008
		(2.49)		(1.00)
SOE		-0.0316		0.0030
		(-0.91)		(0.18)
Top1		0.0022**		0.0011**
		(2.12)		(2.56)
Cons	0.9785***	-7.9156***	0.5176***	-2.2520***
	(348.98)	(-13.84)	(444.07)	(-9.14)
Firm FE	*YES*	*YES*	*YES*	*YES*
Year FE	*YES*	*YES*	*YES*	*YES*
N	37010	37010	37010	37010
adj.R^2	0.73	0.74	0.48	0.49

*Note: The t values are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels respectively. Clustering standard errors are adjusted at the enterprise level. Same as the tables below.*

4.4 Validity Analysis and Robustness Test of DID Estimation

4.4.1 Parallel Trend Test

When employing the difference-in-differences model to evaluate policy effects, the parallel trends assumption is a prerequisite. Specifically, there must be no significant difference in the level of green innovation between enterprises in pilot regions and those in non-pilot regions prior to the implementation of the green finance reform and innovation pilot zone policy. This paper employs the event study method to conduct a parallel trend test, based on established standards in the literature. The results indicate that the regression coefficients for the periods prior to the pilot program were not statistically significant and the 95% confidence intervals included zero. Thus, it can be inferred that the trends in green innovation development were essentially consistent before the pilot, satisfying the prerequisites for the difference-in-differences model. Starting from the period of policy implementation (Period 0), the level of green innovation among enterprises in the pilot zones exhibited a clear upward trend. The coefficients for all periods following implementation were positive and relatively stable, indicating that the green finance policy has a sustained, positive impact on corporate green innovation. It is consistent with the results of the baseline regression analysis in this study.

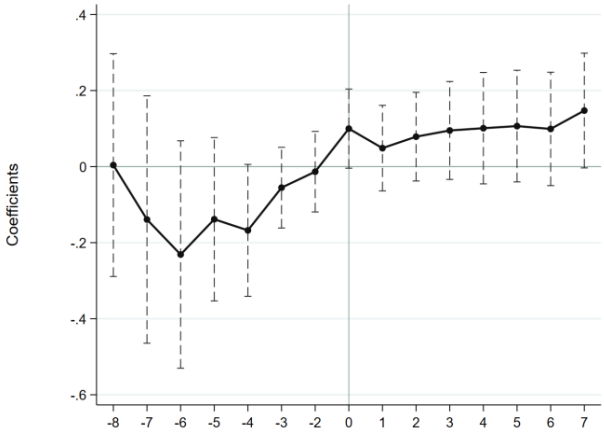
4.4.2 Propensity Score Matching and Entropy Balancing Tests

To mitigate sample self-selection bias, this paper employs the PSM-DID method to conduct robustness tests. Enterprises covered by the pilot zone policies were designated as the treatment group, while the remaining enterprises served as the control group. Covariates selected included firm size, debt-to-asset ratio, return on net assets, operating cash flow, revenue growth rate, years since listing, board size, proportion of independent directors, nature of property rights, and the shareholding ratio of the largest shareholder. Sample matching was conducted using one-to-one nearest neighbor matching, K-nearest neighbor matching, radius

© 2026 The Author(s). Authors retain copyright in their articles.

matching, and kernel matching. The results in Column (1) of Table 5 show that the post-matching DID coefficient is 0.1198, which is significantly positive at the 10% level; the ATT is significantly positive across various matching methods, consistent with the baseline regression findings.

Figure 1: Parallel Trends Test



To circumvent the reliance of traditional matching methods on functional forms and weight specifications, this paper employs the entropy balancing method. By adjusting weights to achieve balance in covariate moment conditions between the treatment and control groups, it mitigates sample selection bias. As shown in Column (2) of Table 5, the DID coefficient after entropy balancing is 0.1249, remaining significantly positive at the 10% level. The sign and magnitude of the coefficient are similar to those of the baseline regression. The above results indicate that, after controlling for differences in observable individual characteristics, the effect of green finance policies in driving corporate green innovation remains robust, confirming the reliability of the baseline findings.

Table 5: Robustness Checks (Propensity Score Matching and Entropy Balancing)

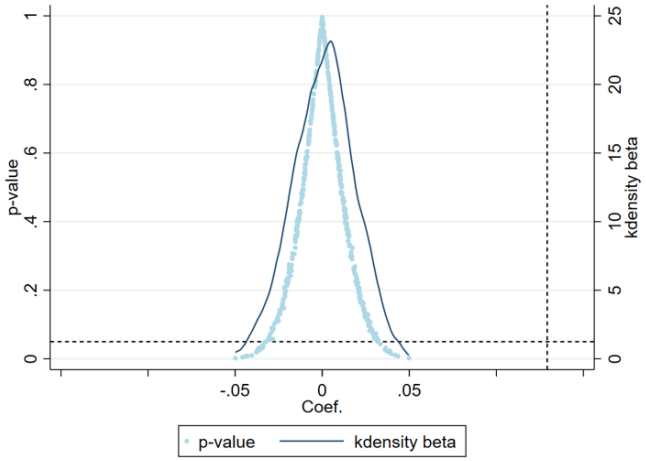
	(1)	(2)
DID	0.1198*	0.1249*
	(1.89)	(1.95)
Size	0.3422***	0.3460***
	(6.07)	(5.97)
Lev	-0.0097	-0.0300
	(-0.05)	(-0.15)
Roe	-0.0650	-0.0626
	(-0.51)	(-0.46)
Cfo	-0.2460	-0.2543
	(-0.91)	(-0.92)
Growth	-0.0321	-0.0242
	(-0.82)	(-0.61)
Age	-0.0637	-0.0701
	(-0.64)	(-0.69)
Board	0.1428	0.1355
	(0.63)	(0.60)
Indep	0.0131**	0.0140**
	(2.15)	(2.28)
SOE	-0.0799	-0.0604
	(-0.67)	(-0.51)
Top1	0.0068*	0.0067*
	(1.94)	(1.86)
Cons	-6.9237***	-6.9533***
	(-3.93)	(-3.88)
Firm FE	*YES*	*YES*
Year FE	*YES*	*YES*
N	3429	3429
adj.R^2	0.72	0.72

© 2026 The Author(s). Authors retain copyright in their articles.

4.4.3 Placebo Test

To rule out the influence of random factors on policy effects and ensure that the baseline results were not obtained by chance, this paper employs a placebo test by randomly selecting a pseudo-treatment group. Keeping the policy implementation timing unchanged, a number of enterprises were randomly selected to form a pseudo-treatment group, and the random sampling and regression processes were repeated 500 times. The results indicate that, after 500 random regressions, most estimated coefficients cluster around zero and follow a normal distribution; furthermore, the majority of p-values exceed 0.05, meaning they are not significant at the 5% level. The baseline regression coefficients and p-valuesdeviate significantly, and far exceed, the levels associated with a random distribution. This indicates that the baseline results cannot be explained by random factors, thereby indirectly demonstrating the robustness of the conclusions.

Figure 2: Placebo Test



4.4.4 Test Using an Alternative Dependent Variable

To avoid estimation bias arising from measurement using a single indicator, this paper conducts robustness checks by replacing the dependent variable with measures based on two dimensions: innovation willingness and innovation quality. The baseline regression measures green innovation using the number of green patent applications. This study re-runs the regression using green innovation intention and patent citation counts as alternative proxies, while keeping the model specification, explanatory variables, control variables, and two-way fixed effects unchanged. Column (1) of Table 6 shows that when the willingness to engage in green innovation is the dependent variable, the DID coefficient is 0.0461, which is significantly positive at the 10% level. In column (2), when the number of patent citations is used as the dependent variable, the DID coefficient is 0.0936, which is significantly positive at the 5% level. Using two alternative indicators, the coefficients of the core variables align in direction and significance with the baseline results, and the results for control variables show no systematic changes. This indicates that the policy incentive effect is significant regardless of whether innovation willingness or innovation quality is measured, and the core conclusions remain robust to the choice of measurement for the dependent variable.

Table 6: Robustness Check (Replacing the Dependent Variable)

	Enterprises' Willingness to Engage in Green Innovation	Corporate Patent Citation Count
	(1)	(2)
DID	0.0461*	0.0936**
	(1.85)	(2.07)
Size	0.1180***	0.0083
	(16.72)	(0.63)
Lev	-0.0161	0.0218
	(-0.63)	(0.54)
Roe	-0.0158	0.0439*
	(-1.06)	(1.90)
Cfo	0.0166	-0.0013

© 2026 The Author(s). Authors retain copyright in their articles.

	Enterprises' Willingness to Engage in Green Innovation	Corporate Patent Citation Count
	(0.43)	(-0.02)
Growth	0.0086	-0.0026
	(1.54)	(-0.34)
Age	-0.0042	-0.0754***
	(-0.28)	(-3.37)
Board	0.0560**	0.0025
	(1.98)	(0.05)
Indep	0.0008	0.0009
	(1.00)	(0.51)
SOE	0.0030	-0.0407
	(0.18)	(-1.12)
Top1	0.0011**	0.0010
	(2.56)	(1.36)
Cons	-2.2520***	0.9673**
	(-9.14)	(2.32)
Firm FE	*YES*	*YES*
Year FE	*YES*	*YES*
N	37010	37010
adj.R2	0.49	0.57

4.4.5 Test by Policy Type

To further enhance the credibility of the baseline findings, this paper categorizes the sample into three time periods based on the policy implementation years of 2017, 2019, and 2022, and employs a three-period difference-in-differences (DID) model to examine the trend of policy effects. Columns (1) through (3) of Table 7 show that the DID coefficient for the first batch of policies (2017) is 0.1279, which is significantly positive at the 10% level. The coefficients for the 2019 and 2022 batches are 0.1694 and 0.1176, respectively, both positive but failing to pass conventional tests of statistical significance. This result is attributable to the broader scope, more comprehensive policy tools, and more robust market response associated with the first batch of pilots, reflecting the more pronounced early effects of the policies. Batch-based testing did not overturn the baseline conclusions but revealed temporal heterogeneity in policy effects, effectively validating the robustness of the baseline results. Among the control variables, firm size is consistently positive and significant at the 1% level; the signs and significance of the other variables remain largely consistent, indicating stable model performance. The adjusted R^2 values range from 0.74 to 0.75, reflecting a good model fit.

4.4.6 Test Excluding Provinces Unaffected by the Policy

To eliminate confounding factors arising from the absence of pilot zones, this paper selects only those provinces that have established pilot zones as the sample for regression analysis. In Column (4) of Table 7, the coefficient of the core explanatory variable (DID) is 0.1405 and is significantly positive at the 5% level. This result is entirely consistent with, and even more significant than, the baseline positive incentive effect, indicating that the policy outcome is not attributable to systematic differences between the groups. The coefficient for the control variable "firm size" is 0.3919 (significant at the 1% level), and the coefficient for the asset-liability ratio is -0.1589 (significant at the 10% level); both align with the results of the baseline regression. The model continues to control for firm and year fixed effects; the adjusted R^2 is 0.73, indicating a high goodness of fit and suggesting that the conclusions are robust.

Table 7: Robustness Tests (Differentiating Policy Types and Excluding Provinces Unaffected by the Policy)

	Impact of the 2017 Policy	Impact of the 2019 Policy	Impact of the 2022 Policy	Excluding Provinces Unaffected by the Policy
	(1)	(2)	(3)	(4)
DID	0.1279*	0.1694	0.1176	0.1405**
	(1.76)	(0.93)	(1.02)	(2.35)
Size	0.3843***	0.3848***	0.3790***	0.3919***
	(19.28)	(18.82)	(18.64)	(12.78)
Lev	-0.0747	-0.0913*	-0.0840	-0.1589*
	(-1.38)	(-1.66)	(-1.52)	(-1.76)
Roe	0.0043	0.0076	0.0078	-0.0729

© 2026 The Author(s). Authors retain copyright in their articles.

	Impact of the 2017 Policy	Impact of the 2019 Policy	Impact of the 2022 Policy	Excluding Provinces Unaffected by the Policy
	(0.14)	(0.25)	(0.26)	(-1.46)
Cfo	-0.0137	-0.0114	0.0037	-0.1069
	(-0.19)	(-0.15)	(0.05)	(-0.88)
Growth	0.0005	0.0017	0.0038	0.0061
	(0.05)	(0.16)	(0.36)	(0.32)
Age	-0.0083	-0.0015	-0.0004	-0.0398
	(-0.31)	(-0.06)	(-0.02)	(-0.85)
Board	0.0891	0.0970	0.1282*	0.2025*
	(1.35)	(1.44)	(1.90)	(1.72)
Indep	0.0044**	0.0040*	0.0047**	0.0066*
	(2.22)	(1.95)	(2.29)	(1.74)
SOE	-0.0312	-0.0313	-0.0313	-0.0430
	(-0.89)	(-0.88)	(-0.88)	(-0.72)
Top1	0.0019*	0.0021*	0.0018*	0.0026
	(1.81)	(1.95)	(1.68)	(1.44)
Cons	-7.8850***	-7.9736***	-7.9411***	-8.0329***
	(-13.64)	(-13.44)	(-13.50)	(-9.27)
Firm FE	*YES*	*YES*	*YES*	*YES*
Year FE	*YES*	*YES*	*YES*	*YES*
N	36340	34754	35064	12611
adj.R^2	0.74	0.75	0.75	0.73

4.4.7 Test of Alternative Fixed Effects

To control for the confounding effects of multidimensional unobservable factors and to test the sensitivity of the findings to various fixed-effects specifications, this paper sequentially introduces fixed effects for province, industry, year-by-industry, and year-by-province into the regression models. Table 8 shows that the DID coefficient is 0.1290 (significant and positive at the 5% level) when province fixed effects are included in column (1), and 0.1257 (significant and positive at the 5% level) when industry fixed effects are included in column (2). When year-industry and year-province interaction fixed effects are included in columns (3) and (4), the DID coefficients are 0.1052 and 0.1515, respectively, both significant and positive at the 10% level. The DID coefficients remain significantly positive across various fixed-effect specifications. Their magnitudes are close to the baseline, and their significance is stable, indicating that the baseline findings are not driven by the choice of fixed effects. The results for the control variables are highly consistent with the baseline. The firm size coefficients are significantly positive at the 1% level, and the adjusted R^2 values range from 0.74 to 0.75, indicating a good model fit. The results remain stable and reliable after controlling for regional, industry, and temporal heterogeneity.

Table 8: Robustness Check (Alternative Fixed Effects)

	(1)	(2)	(3)	(4)
DID	0.1290**	0.1257**	0.1052*	0.1515*
	(2.20)	(2.21)	(1.92)	(1.96)
Size	0.3808***	0.3786***	0.3899***	0.3832***
	(19.94)	(20.01)	(20.74)	(19.96)
Lev	-0.0831	-0.0734	-0.0310	-0.0727
	(-1.55)	(-1.38)	(-0.58)	(-1.35)
Roe	0.0035	-0.0010	-0.0070	0.0078
	(0.12)	(-0.03)	(-0.23)	(0.26)
Cfo	-0.0133	-0.0084	0.0494	-0.0094
	(-0.19)	(-0.12)	(0.70)	(-0.13)
Growth	0.0017	0.0010	-0.0039	0.0010
	(0.16)	(0.10)	(-0.38)	(0.10)
Age	0.0005	-0.0025	0.0448	-0.0111
	(0.02)	(-0.09)	(1.64)	(-0.39)
Board	0.1100*	0.1137*	0.1246**	0.1096*
	(1.72)	(1.76)	(1.98)	(1.70)
Indep	0.0051***	0.0048**	0.0049**	0.0054***

© 2026 The Author(s). Authors retain copyright in their articles.

	(1)	(2)	(3)	(4)
	(2.64)	(2.48)	(2.57)	(2.78)
SOE	-0.0220	-0.0329	-0.0257	-0.0194
	(-0.63)	(-0.95)	(-0.73)	(-0.55)
Top1	0.0023**	0.0022**	0.0015	0.0020*
	(2.26)	(2.13)	(1.44)	(1.92)
Cons	-7.9973***	-7.9059***	-8.7342***	-7.9170***
	(-14.23)	(-14.19)	(-15.66)	(-13.77)
Firm FE	*YES*	*YES*	*YES*	*YES*
Year FE	*YES*	*YES*	*NO*	*NO*
Province FE	*YES*	*NO*	*NO*	*NO*
Industry FE	*NO*	*YES*	*NO*	*NO*
Year×Industry FE	*NO*	*NO*	*YES*	*NO*
Year×Province FE	*NO*	*NO*	*NO*	*YES*
N	37010	37010	36997	37008
adj.R^2	0.74	0.75	0.75	0.74

5. Cross-sectional Difference Analysis

The impact of policies establishing pilot zones for green finance reform and innovation on corporate green innovation exhibits significant cross-sectional divergence at the firm level. A company's internal governance foundation and its basis for environmental information disclosure directly moderate the strength of the policy incentives. This section analyzes the subject from two dimensions, the substitutive role of corporate governance and the spillover effects of information disclosure, to clarify the logic behind the differentiated impact of policy incentives on enterprises with varying endowments.

5.1 The Substitutive Role of Corporate Governance

Internal controls and a standardized environmental management system together form a comprehensive internal green governance system. The two complement and check each other, establishing an internal mechanism of constraint.

On one hand, the effectiveness of internal control constitutes a crucial foundation for a firm's internal governance, directly influencing the management of R&D funds, the constraint of agency costs, and the soundness of innovation decision-making. For enterprises with deficiencies in internal controls, the use of R&D funds lacks constraints, which hinders the pursuit of green innovation. This paper uses the effectiveness of internal control (IsValid) as a measure of a firm's internal governance quality, assigning a value of 1 if internal control is effective and 0 otherwise, and incorporates an interaction term between this variable and the difference-in-differences (DID) term into the baseline model for empirical testing. The results in Column (1) of Table 9 show that the baseline DID coefficient is significantly positive at the 1% level, indicating that green innovation in firms with weak internal controls significantly increased following the policy shock. The coefficient of the interaction term between policy and internal control effectiveness is significantly negative, demonstrating that green finance policies exert a stronger incentive effect on green innovation for enterprises with weak internal controls.

On the other hand, ISO 14001 environmental management system certification reflects a company's specialized environmental governance capabilities. It complements internal controls and further validates the logic of governance substitution. Enterprises that have obtained this certification possess mature mechanisms for environmental risk prevention and control as well as green operations, with stable institutional support for green innovation. In contrast, enterprises that have not yet obtained the certification tend to have fragmented environmental management and lack institutional safeguards for green R&D. This paper uses ISO14001 certification status (IsPassISO14001) to measure the level of corporate environmental governance; a value of 1 is assigned for certification and 0 otherwise. Firms scoring above the annual median are classified into the high internal governance group (IsPassISO14001 = 1), while the remainder are assigned to the low internal governance group (IsPassISO14001 = 0). The results in Column (2) of Table 9 show that the baseline DID coefficient is significantly positive at the 5% level, indicating that enterprises with imperfect environmental management systems experienced a significant improvement in their green innovation performance driven by the policy. The coefficient of the interaction term between policy and certification is significantly negative,

© 2026 The Author(s). Authors retain copyright in their articles.

indicating that the marginal incentive effect of the policy on enterprises with mature environmental management systems is relatively limited.

Synthesizing the empirical findings regarding internal control effectiveness and ISO 14001 environmental management certification, it is evident that for enterprises suffering from internal control failures or lacking standardized environmental management systems, external policy constraints can compensate for the absence of internal governance mechanisms, thereby generating stronger incentives for green innovation. In contrast, for enterprises with robust internal governance and comprehensive environmental control systems, which already possess mature operational mechanisms for innovation, the incremental gains resulting from such policies are relatively limited.

Table 9: Cross-sectional Regression Results of Corporate Governance Substitution Effect

	(1)	(2)
Did	0.6120***	0.1498**
	(2.99)	(2.52)
IsValid	0.1558***	
	(3.53)	
c.Did#c.IsValid	-0.4806**	
	(-2.42)	
IsPassISO14001t		0.0116
		(0.86)
c.DID#c.IsPassISO14001		-0.1005**
		(-2.10)
Size	0.3774***	0.3803***
	(19.53)	(19.28)
Lev	-0.0750	-0.0797
	(-1.36)	(-1.47)
Roe	-0.0114	-0.0003
	(-0.37)	(-0.01)
Cfo	-0.0375	-0.0176
	(-0.52)	(-0.24)
Growth	0.0012	0.0018
	(0.11)	(0.17)
Age	-0.0018	-0.0044
	(-0.07)	(-0.16)
Board	0.1110*	0.1142*
	(1.74)	(1.74)
Indep	0.0045**	0.0049**
	(2.38)	(2.50)
SOE	-0.0343	-0.0318
	(-0.98)	(-0.92)
Top1	0.0021**	0.0022**
	(2.05)	(2.11)
Cons	-8.0151***	-7.9241***
	(-14.27)	(-13.85)
Firm FE	*YES*	*YES*
Year FE	*YES*	*YES*
N	36091	37010
adj. R^2	0.74	0.74

5.2 Extended Effect of Information Disclosure

Corporate environmental information disclosure encompasses the disclosure of environmental philosophies at the fundamental strategic level and the disclosure of social responsibility reports at the comprehensive operational level. The quality of these two types of disclosure collectively determines the enterprise's endogenous drive for green transformation and the transparency of its information to external stakeholders.

From the perspective of strategic disclosure, companies that proactively disclose their environmental policies and green development models have integrated environmental goals into their daily operations, with green innovation driven by internal momentum. Conversely, companies that do not disclose their

© 2026 The Author(s). Authors retain copyright in their articles.

environmental philosophies show a weaker willingness to undergo a green transition, relying more on external policy drivers for green R&D. This study employs the disclosure of environmental protection concepts (EPtConcept) to measure the strategic orientation of corporate environmental protection; a value of 1 is assigned if a company discloses information such as environmental protection concepts, environmental policies, or green development models (0 otherwise), and the sample is divided into high and low groups based on the annual median for analysis. The results in Column (1) of Table 10 show that the baseline DID coefficient is significantly positive at the 5% level, indicating a significant increase in green innovation among enterprises that did not disclose environmental philosophies following the policy's implementation. The coefficient for the interaction term between the policy and environmental philosophy disclosure is significantly negative, suggesting that the policy has a stronger effect on boosting innovation in enterprises with insufficient disclosure of environmental philosophies.

From the perspective of integrated business operations, enterprises that consistently provide high-quality social responsibility disclosures demonstrate greater information transparency and find it easier to secure credit for green projects. Conversely, enterprises with low levels of disclosure fail to adequately communicate green-related information to the outside world and face tighter financial constraints regarding green R&D. This paper uses the disclosure level of corporate social responsibility (CSR) reports to measure the extent of corporate information disclosure. Based on the upper quartile of disclosure quality for the sample years, firms with disclosure levels exceeding the annual upper quartile are classified into the high-disclosure group (IsPublishCSRRep = 1), while the remainder are classified into the low-disclosure group (IsPublishCSRRep = 0). When the grouping variable and its interaction term with the DID variable are incorporated into the baseline model to examine the information disclosure mechanism, the results in Column (2) of Table 10 show that the baseline DID coefficient is significantly positive at the 5% level, indicating that green innovation among firms with low disclosure levels increased significantly following the policy intervention. Meanwhile, the coefficient for the interaction term between the policy and the high-disclosure group is significantly negative, implying that the policy's marginal effect on promoting innovation among firms with high information disclosure is relatively limited.

By synthesizing the empirical findings regarding environmental philosophy disclosure and social responsibility reporting, it can be concluded that policies aimed at addressing information disclosure deficiencies are more effective for enterprises that previously lacked environmental philosophy disclosures, driving greater green innovation. Conversely, for enterprises that already possess robust information disclosure mechanisms, the scope for incremental improvement resulting from such policies is more limited, given the already low level of information asymmetry.

Table 10: Cross-sectional Regression Results for the Spillover Effect of Information Disclosure

	(1)	(2)
DID	0.1429**	0.1393**
	(2.41)	(2.36)
EPtConcept	-0.0153	
	(-1.12)	
c.DID#c.EPtConcep	-0.0984*	
	(-1.81)	
IsPublishCSRReport		0.1064***
		(4.50)
c.did#c.IsPublishCSRReport		-0.2122**
		(-2.51)
Size	0.3806***	0.3802***
	(19.32)	(19.31)
Lev	-0.0788	-0.0843
	(-1.45)	(-1.56)
Roe	-0.0005	-0.0000
	(-0.02)	(-0.00)
Cfo	-0.0180	-0.0200
	(-0.25)	(-0.28)
Growth	0.0016	0.0021
	(0.15)	(0.20)
Age	-0.0045	-0.0055

© 2026 The Author(s). Authors retain copyright in their articles.

	(1)	(2)
	(-0.17)	(-0.21)
Board	0.1152*	0.1159*
	(1.75)	(1.76)
Indep	0.0049**	0.0049**
	(2.49)	(2.52)
SOE	-0.0311	-0.0341
	(-0.90)	(-0.99)
Top1	0.0022**	0.0023**
	(2.09)	(2.18)
Cons	-7.9279***	-7.9173***
	(-13.86)	(-13.91)
Firm FE	*YES*	*YES*
Year FE	*YES*	*YES*
N	37010	37010
adj. R^2	0.74	0.74

6. Conclusions and Policy Recommendations

Green finance reform and innovation pilot zones represent a significant institutional exploration of China's financial supply-side structural reform within the realm of green development. The policy outcomes of these zones not only determine the practical success of local green transitions but also shape the strategic path for the financial system to empower the high-quality development of the real economy in the context of the "dual carbon" goals.

Taking Shanghai and Shenzhen A-share listed companies from 2014 to 2024 as the research subjects, and leveraging the quasi-natural experiment created by the phased establishment of green finance reform and innovation pilot zones, this paper constructs a multi-period difference-in-differences (DID) model to systematically evaluate the policy effects from the perspectives of innovation intensity and innovation willingness. Research indicates that policies in the pilot zones can significantly boost the level of corporate green innovation; the volume of green patent applications by pilot-zone enterprises is approximately 12.92% higher than that of enterprises in non-pilot areas. The policy effects are more pronounced in companies characterized by weak internal controls, a lack of environmental management mechanisms, low transparency regarding environmental philosophies, and poor-quality social responsibility reports. Meanwhile, in companies with higher asset operational efficiency, the policy dividends were more effectively translated into tangible R&D outcomes.

Based on the above findings, the following policy recommendations are proposed:

It is recommended to systematically scale up the experience gained in pilot zones and refine the green finance framework based on "unified standards plus local characteristics." Elevate core rules, such as the certification of proven green projects and environmental credit assessments, from the pilot zones to national standards to reduce cross-regional institutional friction. At the same time, allow space for local innovation by enabling the design of tailored toolkits based on distinct industrial structures and resource endowments. This fosters deep alignment between green finance, the cultivation of "new quality productive forces," and "dual carbon" goals, while channeling medium- and long-term capital toward key areas of green technology breakthroughs.

It is also recommended to strengthen external governance and information transparency, and to implement targeted support based on classification. On one hand, it is necessary to refine the differentiated credit mechanism that rewards excellence and penalizes poor performance by incorporating both environmental compliance and innovation incentives into credit assessments. This involves formulating unified detailed rules for environmental information disclosure by listed companies, establishing an inter-departmental information-sharing platform, and reducing the costs associated with screening green projects for financing. On the other hand, for enterprises lagging in their transition, efforts should be intensified to lower entry barriers while applying pressure to compel change, accompanied by the provision of transition finance instruments and compliance guidance services. Meanwhile, high-quality enterprises should receive robust support to enhance

© 2026 The Author(s). Authors retain copyright in their articles.

their capabilities, enabling them to secure long-term, low-interest capital through mechanisms such as green bonds and the pledging of carbon assets.

References

[1] Zhang, W. Q., & Chen, Y. K. (2022). Research on the impact of environmental regulation tools and R&D subsidies on green technology innovation. Ecological Economy, 38(1), 36–46.

[2] Chen, W., Lin, R., & Ma, G. (2026). Responsibility in the shadows: How does firms' shadow banking reshape carbon emission performance? Finance Research Letters, Article 110163. https://doi.org/10.1016/J.FRL.2026.110163

[3] Xie, W. Y., Li, M. X., & Yang, K. (2025). Research on the mechanism of patient capital on the green transformation and development of heavily polluting enterprises. Financial Theory & Practice, (11), 20–30.

[4] Li, S. H., You, B. Y., Niu, K., et al. (2025). Green credit, informal financial financing and green innovation of heavily polluting enterprises. Economic Research Journal, 60(9), 198–217.

[5] Sun, X. Y., & Shao, J. J. (2025). Research on the impact of green credit policy on the quality of enterprise green innovation. Time Economy and Trade, 22(6), 61–66. https://doi.org/10.19463/j.cnki.sdjm.2025.06.008

[6] Ma, Z. H., Gu, H. Y., & Wang, Y. X. (2026). Government guidance funds, institutional friction and enterprise innovation quality: From the perspective of patient capital. Studies in Science of Science, 1–14. Advance online publication. https://doi.org/10.16192/j.cnki.1003-2053.20260717.002

[7] Chen, M., Xu, H., Tian, F., et al. (2026). The impact of China's R&D and innovation strategy on total factor productivity of listed intelligent manufacturing firms. Sustainability, 18(10), 5128. https://doi.org/10.3390/SU18105128

[8] Cui, H., Lu, Z., & Li, H. M. (2026). Green finance reform and innovation pilot zone policies and enterprise green productivity. Research on Economics and Management, 47(7), 62–77. https://doi.org/10.13502/j.cnki.issn1000-7636.2026.07.005

[9] Wang, Y., Wang, J. Z., Wu, Z. S., et al. (2025). Can green finance promote the economy's shift from virtual to real? Empirical evidence based on green finance reform and innovation pilot zones. Journal of Financial Research, (12), 20–38.

[10] Lü, K. J., & Zhu, X. G. (2026). Green finance and clean energy industry: Evidence from the scale of photovoltaic power stations under regional development differentiation. Journal of Shanghai University (Social Sciences Edition), 43(4), 132–148.

[11] Zhang, R., & Ling, S. (2026). The synergistic effect of environmental tax and green finance policy on corporate green technology innovation: Empirical evidence from Chinese listed firms. Sustainability, 18(9), 4502. https://doi.org/10.3390/SU18094502

[12] Zhou, Y. F. (2025). Research on green technology innovation empowering high-quality regional economic development from the perspective of innovation factor recombination and institutional synergy. Jiangsu Science & Technology Information, 42(20), 16–22.

[13] Zeng, H. X., & Ding, N. (2026). Research on the composite effect of equity incentives and employee stock ownership on enterprise green innovation. Business Accounting, (13), 4–10.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

© 2026 The Author(s). Authors retain copyright in their articles.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.1832
Published by: Zeus Press

Green Finance and Corporate Green Innovation

Xinyu Wen*

School of Economics and Management, Beijing Forestry University, Beijing 100083, China

**Corresponding author: Xinyu Wen.*

Abstract

Green finance has drawn considerable attention from both academia and policymakers as a key policy instrument for promoting green development and economic transformation in China. Against the backdrop of the central government's strategic efforts to establish a green financial system and a market-oriented green technology innovation system, this paper takes the Pilot Zones for Green Finance Reform and Innovations policy as a point of entry to examine its impact on corporate green innovation. The findings indicate that the pilot zone policy significantly enhances the level of green innovation among enterprises within the pilot regions. Further analysis reveals that the policy effect is more pronounced among firms with lower management ownership, weaker among those with larger pre-existing government green subsidies, and stronger among firms subject to higher levels of media coverage. This study provides empirical evidence for understanding green finance policies. Advancing the Pilot Zones for Green Finance Reform and Innovations policy to drive corporate green innovation is essential for promoting high-quality economic development.

Keywords

green finance, pilot zones for green finance reform and innovations, green innovation, staggered difference-in-differences

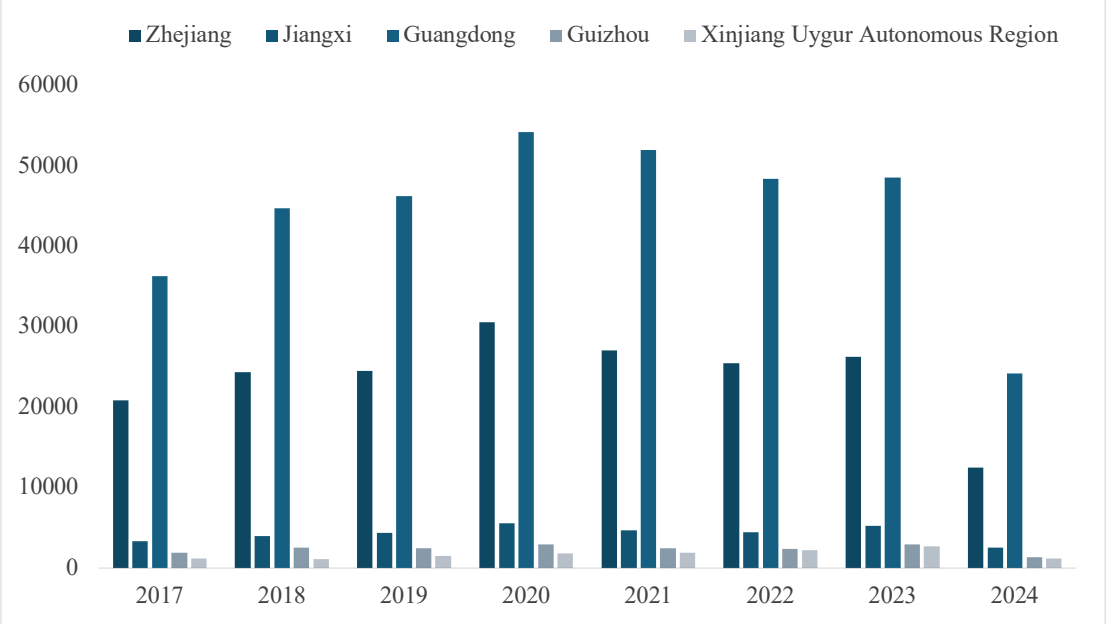
1. Introduction

The report to the 20th National Congress of the Communist Party of China explicitly states that "advancing the green and low-carbon transformation of economic and social development is a key link in achieving high-quality development." Establishing Pilot Zones for Green Finance Reform and Innovation is crucial to realizing the "dual carbon" goals and holds great significance for driving sustainable economic growth and fostering modernization characterized by the harmonious coexistence of humanity and nature. To establish a differentiated policy framework covering areas such as green credit, green bonds, and environmental information disclosure, seven government agencies, including the People's Bank of China, the National Development and Reform Commission, and the Ministry of Finance, have, since 2017, designated specific provinces and regions as pilot zones for green finance reform and innovation, rolling out the initiative in three batches. At present, research on the effects of this policy has proceeded from a macro-regional perspective, revealing that the pilot zone policy can foster the ecological transformation of industrial structures and empower the development of "new quality productive forces" in cities [1, 2]. On the other hand, examining the issue from the perspective of micro-level corporate behavior reveals that the policy can enhance corporate ESG performance and drive a shift in corporate activity from the financial sector to the real economy [3, 4].

© 2026 The Author(s). Authors retain copyright in their articles.

As one of the first Pilot Zones for Green Finance Reform and Innovation in China, Huzhou has, from its approval in 2017 through the end of the second quarter of 2025, achieved a total of 74 green finance reform and innovation outcomes that were the first of their kind nationwide. Over the eight years since the pilot zone was established, green loans across the city have grown at an average annual rate of 41%. By the end of the second quarter of 2025, green loans accounted for 33.8% of total loans, far exceeding the national and provincial averages. Take Zhejiang Yikeou Environmental Protection Technology Co., Ltd., located within the pilot zone, as an example: backed by green credit support from the Bank of Huzhou, the company was included in the National Intellectual Property Administration's inaugural list of patent-intensive products in 2024; its "Engineering Laboratory for Pyrolysis Treatment and Resource Utilization of Oily Solid Waste in the Petroleum and Chemical Industry" was certified by the China Petroleum and Chemical Industry Federation in 2023; it was selected as a key "Little Giant" enterprise (specialized, refined, distinctive, and innovative) at the national level in 2021; and in 2020, it won the First Prize for Scientific and Technological Progress in Green Manufacturing from the Chinese Mechanical Engineering Society. Taking the five provinces (regions) covered by the first batch of pilot zones as an example, from the establishment of the pilot zone in 2017 to the end of 2024, the cumulative number of green patent applications in Guangdong Province reached 355,000, Zhejiang Province reached 192,000, Jiangxi Province reached 34,000, Guizhou Province reached 19,000, and Xinjiang Uygur Autonomous Region reached 14,000. The total number of applications in the five provinces (regions) in eight years has exceeded 610,000. Based on this, this article takes Shanghai and Shenzhen A-share non-financial listed companies from 2011 to 2024 as a sample to study the effect, heterogeneity conditions and internal mechanisms of green finance reform and innovation pilot zone policies on corporate green innovation.

Figure 1: Number of Green Patent Applications in the Five Pilot Provinces (Regions)



Note: Data sourced from the National Intellectual Property Administration and CNRDS.

This paper first employs a multi-period difference-in-differences model, using the volume of corporate green patent applications as a measure of green innovation performance. After controlling for firm and year fixed effects, the study finds that the pilot zone policy significantly boosts corporate green innovation output. This paper verifies the reliability of the baseline results, and confirms the robustness of the conclusions under various conditions, using multiple methods, including parallel trend tests, placebo tests, propensity score matching, and entropy balancing. Furthermore, this paper examines the sources of heterogeneity in policy effects, finding that the proportion of management shareholding amplifies the policy's innovation-incentivizing effect by mitigating agency conflicts. Pre-existing stocks of government green subsidies diminish the marginal impact of new policies due to resource lock-in; and media attention magnifies the policy's influence by alleviating information asymmetry and strengthening reputational constraints. Finally, the paper offers specific recommendations for optimizing the mix of green finance policy instruments and improving supporting information disclosure mechanisms.

© 2026 The Author(s). Authors retain copyright in their articles.

The research contributions of this paper are primarily reflected in the following three aspects. First, this paper expands the research boundaries regarding the economic consequences of green finance. Existing literature largely evaluates green finance policies from the perspectives of macro-regional carbon emissions or aggregate corporate financing constraints [5]. In contrast, this paper focuses on green technological innovation at the firm level and deepens the understanding of the mechanisms through which Pilot Zones for Green Finance Reform and Innovation exert their influence, adopting a tripartite perspective that encompasses internal corporate governance, government resource allocation, and the external information environment.

Secondly, this paper enriches the research on the relationship between industrial policy and corporate innovation. Previous literature on how industrial policy affects corporate innovation has largely focused on direct fiscal subsidies or tax incentives [6, 7], with relatively little discussion regarding the impact of industrial policy in the form of financial market reforms. This paper finds that the green financial reform and innovation pilot zone policy can significantly affect the green innovation behavior of enterprises, supplementing the channels through which industrial policies affect corporate innovation.

Finally, this article has strong practical guidance significance. On one hand, the findings regarding heterogeneity suggest that regulators, when expanding green finance pilot zones, should consider the policy's suitability for enterprises with varying levels of governance and resource endowments, and leverage external information sources, such as the media, to amplify the policy's positive guiding effect. On the other hand, this paper identifies dimensions for market investors to assess corporate green innovation capabilities and offers a novel perspective on the effective alignment between capital markets and industrial policy.

The remainder of the paper is organized as follows: Section 2 outlines the institutional background; Section 3 details the research design; Section 4 presents the main empirical results; Section 5 conducts heterogeneity analysis; and the final section summarizes the study and offers policy recommendations.

2. Institutional Background

In September 2015, the CPC Central Committee and the State Council issued the Overall Plan for the Reform of the Ecological Civilization System, proposing for the first time the establishment of a green financial system. In August 2016, seven government agencies, including the People's Bank of China, the National Development and Reform Commission, the Ministry of Finance, the Ministry of Environmental Protection, and the regulatory commissions for banking, securities, and insurance, issued the Guiding Opinions on Establishing a Green Financial System, which defined green finance and outlined the policy framework. In June 2017, an executive meeting of the State Council decided to establish Pilot Zones for Green Finance Reform and Innovation in five provinces and autonomous regions: Zhejiang, Jiangxi, Guangdong, Guizhou, and Xinjiang. Seven government departments later jointly issued master plans for five provinces (and regions), designating the Ganjiang New Area in Jiangxi, the Gui'an New Area in Guizhou, Huzhou and Quzhou in Zhejiang, the Huadu District of Guangzhou in Guangdong, and the cities of Hami, Karamay, and Changji Prefecture in Xinjiang as the first batch of pilot zones. These zones aim to explore pathways for regional green transformation through financial mechanism innovation, accumulate replicable experience for the nation's ecological civilization development, and foster the coordinated and sustainable development of regional economies and the ecological environment. In November 2019, Lanzhou New District of Gansu Province was approved to join, further strengthening the team of the pilot area, helping to verify the effectiveness of green finance policies on a wider scale, and injecting new financial vitality into the green revitalization and ecological protection of the western region. In August 2022, Chongqing City was approved for establishment, which has important demonstration significance in promoting the green and high-quality development of the Yangtze River Economic Belt and realizing the modernization of harmonious coexistence between man and nature. As of the end of 2024, the pilot zone covered ten locations across seven provinces (and regions). In April 2024, seven ministries and commissions issued the Opinions on Further Strengthening the Construction of Green Financial Reform and Innovation Pilot Zones and proposed 21 specific measures, marking the policy's shift from pilot exploration to in-depth promotion, which will help promote the comprehensive, coordinated and sustainable development of the economy and society.

Innovative products and services are one of the core tasks of the overall plan of each pilot zone. The Implementation Rules for Building a Green Finance Reform and Innovation Pilot Zone in Ganjiang New Area outline 46 tasks across 14 areas. The rules place special emphasis on fostering green industrial development

© 2026 The Author(s). Authors retain copyright in their articles.

by supporting energy-saving and consumption-reducing technical upgrades for enterprises in sectors such as iron and steel, non-ferrous metals, papermaking, and printing and dyeing, while accelerating the cultivation of green products and demonstration projects for green factories. Key priorities include supporting the listing and public trading of green enterprises and establishing green industry development funds, thereby using financial instruments to incentivize companies to implement source-based pollution control and process innovation. Huzhou City released the Catalogue of Financial Support Economic Activities for Textile Industry Transformation in 2022, covering 106 low-carbon transformation technology paths in 37 sub-sectors in 9 categories. It includes "energy-saving technical transformation equipment update", "energy efficiency improvement" and "clean energy substitution" into the scope of support, and guide enterprises to invest credit resources in the research and development and transformation of green production processes. The Implementation Rules for Building Pilot Zones for Green Finance Reform and Innovation in Guangzhou, Guangdong Province explicitly propose "expanding the scope of eligible collateral and pledges for green enterprises and projects." By exploring the use of environmental rights, such as carbon emission allowances, pollutant discharge rights, and energy consumption rights, along with their associated revenue rights, as qualified collateral, the rules open up new financing channels for green technology enterprises, effectively resolving the challenge of securing financing for asset-light green innovation companies. Quzhou City has established a "1+6+N" policy framework centered on carbon account finance and introduced digital tools, such as the "Carbon Emissions e-Ledger" and "Carbon Credit e-Report", to provide targeted financing support for green technology R&D. Gui'an New Area has allocated special incentive funds to provide rewards and subsidies to financial institutions that develop green credit, green bonds, and green insurance. In 2019, the People's Bank of China issued a document supporting enterprises within the pilot zones in issuing green debt financing instruments, establishing "green channels," and encouraging product innovation. By employing market-based mechanisms to alleviate financing constraints on corporate green innovation and reduce the R&D risks associated with green products and services, the policy demonstrates China's strong determination to reform and unleash the vitality of enterprises as key drivers of green innovation.

Existing literature has evaluated the economic consequences of policies regarding Pilot Zones for Green Finance Reform and Innovation. At the macro level, existing research indicates that policies in pilot zones effectively facilitate the eco-friendly transformation of industrial structures and drive the real economy to shift away from "fictitious" activities and toward the real sector [1, 4]. At the micro level, existing literature confirms that the establishment of pilot zones can improve corporate ESG performance [3]. Research on the drivers of green innovation has long examined the subject from the perspectives of industrial policy and macroeconomic regulation. Dong Zhaoyingzi et al. found that energy-saving and emission-reduction policies can promote both the quantity and quality of green innovation [8]. Liu Jinke et al. found that environmental protection taxes can significantly enhance corporate green innovation capabilities [9]. Yu Zhimai found that government environmental subsidies exert a significant negative moderating effect on corporate green innovation [10]. Furthermore, from the perspective of the transmission of environmental practices among firms, Yu xinguo found that the greater the pressure of environmental competition from peers, the more inclined firms are to engage in green innovation [11]. Thus, existing research focuses, on the one hand, on the impact of policies regarding green finance reform and innovation pilot zones on aspects such as corporate environmental investment, ESG performance, and the alleviation of financing constraints; on the other hand, it largely explores the drivers of corporate green innovation from perspectives such as environmental regulation, government subsidies, and corporate strategy. Few studies have analyzed the two in an integrated manner, creating an opportunity for this paper to explore the effects and boundary conditions of policies establishing green finance reform and innovation pilot zones on corporate green innovation.

Policies for Pilot Zones for Green Finance Reform and Innovation internalize environmental externalities through market-based mechanisms, guiding the allocation of financial resources toward green innovation. Existing research perspectives on green finance policies and corporate green innovation can be broadly categorized into two types. Some scholars believe that green finance policies promote green innovation. Wang Xin and Wang Ying point out that green credit policies significantly promote green innovation in industries subject to green credit restrictions by reducing agency costs and enhancing investment efficiency [5]. Lu Jing and others found that green credit policies function by fostering technological innovation among polluting enterprises, driving the reallocation of credit resources from polluting industries to clean industries, and thereby achieving optimal resource allocation at the industry level [12]. Policies regarding Pilot Zones for Green Finance Reform and Innovation are regional in nature. They further integrate diverse financial

© 2026 The Author(s). Authors retain copyright in their articles.

instruments, such as green bonds, green insurance, and pledges or collateral based on environmental rights and interests, and grant localities greater scope for institutional innovation, thereby providing an opportunity to re-examine the innovation effects of green finance.

However, some scholars have also expressed cautious views regarding macro-industrial policies such as green finance. Research by Li Shihui and others indicates that credit constraints leave enterprises in a predicament where they lack the means to act despite the willingness to do so, resulting in a crowding-out effect on their green innovation [13]. Li Wenjing and Zheng Manni point out that macroeconomic industrial policies may incline enterprises to engage in strategic innovations, such as utility model patents, which are less demanding, to cater to regulatory requirements, rather than pursuing substantive green invention patent innovations [6]. Guo Junjie et al. found that when green credit policies lack sufficient tolerance for short-term innovation failure, they inhibit innovation through crowding-out effects and resource constraints [14]. Wang Shiqi and Wang Yongqiao found that green innovation induced by green credit policies acts as a leverage effect superimposed on existing innovation activities. While it does not crowd out other forms of technological innovation, it is essentially a continuation of existing innovation paths rather than a breakthrough [15].

This paper argues that the positive impact of pilot zone policies on corporate green innovation is dominant. Policies in the pilot zone require financial institutions to innovate green credit products and broaden the scope of collateral and pledges, incorporating assets such as carbon emission rights, pollutant discharge rights, and energy-use rights, thereby lowering financing thresholds and costs for green innovation projects. Policies in the pilot zone leverage market-based mechanisms, such as green credit incentives and environmental regulatory constraints, to compel enterprises to improve their internal governance. Adjusting internal governance processes in this way is more conducive to sustained research and development of green technologies. By incorporating corporate environmental information disclosure into credit rating systems, the policy subjects enterprises to reputational consequences and the risk of penalties for non-compliance. This compels them to optimize their internal governance structures, thereby enhancing their level of green innovation. Furthermore, under the traditional policy framework, certain enterprises, due to factors such as scale, ownership structure, or ties between government and business, find it easier to secure government subsidies, thereby creating implicit barriers. Policies in the pilot zones have reduced corporate reliance on direct government subsidies through market-based financing mechanisms, providing more equal green financing opportunities for small and medium-sized enterprises (SMEs) and non-state-owned enterprises. Media coverage extends the reach of policy transmission and alleviates information asymmetry between enterprises and financial institutions, ensuring that policy signals more effectively reach corporate decision-makers. Based on this, this paper proposes the following hypotheses:

H1: The implementation of policies for green finance reform and innovation pilot zones can significantly promote corporate green innovation.

3. Research Design

3.1 SAMPLE Selection and Data Sources

This study uses Shanghai and Shenzhen A-share listed companies from 2011 to 2024 as the initial research sample. Pilot Zones for Green Finance Reform and Innovation were established in three batches, in 2017, 2019, and 2022, ensuring the sample period fully covers the stages before and after policy implementation. To ensure the representativeness of the sample, the following screening procedures were applied: (1) excluding firms in the finance and insurance sectors; (2) excluding firms with abnormal trading statuses (such as ST and *ST) and samples with severe missing data for key variables; and (3) winsorizing all continuous variables at the 1st and 99th percentiles. Financial and green patent data for listed companies were obtained from the CSMAR database, while the list of enterprises in the pilot zones and the timing of relevant policies were manually compiled based on official documents released by the People’s Bank of China and other government agencies.

3.2 Model Design and Variable Selection

To examine the impact of the construction of Pilot Zones for Green Finance Reform and Innovation on the green transformation of enterprises, this paper constructs the following multi-period difference-in-differences model (1):

© 2026 The Author(s). Authors retain copyright in their articles.

$$Patent_{i,t}=\alpha+\beta\cdot Did_{i,t}+\gamma\cdot Controls_{i,t}+\mu_i+\lambda_t+\varepsilon_{i,t} \quad (1)$$

The main variables included in Model (1) are as follows:

3.2.1 Dependent Variable: Corporate Green Transformation (Patent)

Drawing on the work of Wang Yao et al. and Li Wenjing et al., this paper uses the number of corporate green patent applications to measure the degree of green transformation [4, 6]. Compared to the number of granted patents, the volume of patent applications reflects enterprises' willingness to innovate more promptly and is less affected by the examination cycle. Given that the distribution of this variable is right-skewed, the natural logarithm was taken after adding 1 to the raw application count.

3.2.2 Core Explanatory Variable: Multi-period DID (DID)

This paper treats the establishment of green finance reform and innovation pilot zones as a quasi-natural experiment and employs a multi-period difference-in-differences (DID) method. The variable is assigned values based on the specific year (2017, 2019, or 2022) in which the pilot zone where the enterprise is located was approved: it takes a value of 1 if the enterprise is registered in a pilot zone and the time period is the year of policy implementation or later, and 0 otherwise.

3.2.3 Control Variables

In order to alleviate the omitted variable bias, this article selects two types of control variables: the first type is enterprise characteristic variables, including asset-liability ratio (*Lev*), return on net assets (*Roe*), operating cash flow (*Cfo*), growth (*Growth*), enterprise size (*Size*) and enterprise age (*Age*); the second category is corporate governance variables, including ownership concentration (*Top1*), board size (*Board*), proportion of independent directors (*Indep*) and nature of property rights (*SOE*).

Table 1: Description of Key Variables

Variable Type	Variable Name	Symbol	Definition and Measurement
Dependent Variable	Green Patent Applications	*Patent*	Ln(1 + number of green patent applications)
Core Explanatory Variable	Multi-period DID	*Did*	1 if the firm's registered location is in a pilot zone and the year is the policy implementation year or later; 0 otherwise
Control Variables	Firm Size	*Size*	Ln(total assets)
	Leverage	*Lev*	Total liabilities / total assets
	Return on Equity	*Roe*	Net profit / net assets
	Operating Cash Flow	*Cfo*	Net cash flow from operating activities / total assets
	Growth	*Growth*	Operating revenue growth rate
	Firm Age	*Age*	Ln(current year − listing year + 1)
	Board Size	*Board*	Ln(number of board members)
	Independent Director Ratio	*Indep*	Number of independent directors / total number of board members × 100
	Ownership Nature	*SOE*	1 if state-owned enterprise; 0 if non-state-owned enterprise
	Largest Shareholder Ownership	*Top1*	Shares held by the largest shareholder / total shares × 100

4. Empirical Results and Analysis

4.1 Descriptive Statistics

Table 2 reports the descriptive statistics for the main variables. The mean value of green patent applications is 0.920, with a standard deviation of 1.204, indicating that the overall level of green innovation among the sample firms is relatively low and that there are significant disparities between them. The mean value of the multi-period DID variable is 0.040, indicating that approximately 4% of the observations in the full sample fall into the treatment group category following policy implementation. This aligns with the policy context of the green finance reform and innovation pilot zones being rolled out in three batches across 2017, 2019, and 2022. Regarding the control variables, the mean values are 22.273 for firm size (*Size*), 0.427 for the asset-liability ratio (*Lev*), 0.044 for return on equity (*Roe*), and 0.339 for ownership nature (*SOE*). The distributions

© 2026 The Author(s). Authors retain copyright in their articles.

of these variables fall within reasonable ranges, with no extreme outliers, indicating that the sample data quality is high and suitable for subsequent regression analysis.

Table 2: Descriptive Statistics

	N	*Mean*	*SD*	*Min*	*p25*	*Median*	*p75*	*Max*
Patent	43182	0.920	1.204	0.000	0.000	0.000	1.609	4.787
Did	43182	0.040	0.196	0.000	0.000	0.000	0.000	1.000
Size	43182	22.273	1.303	19.141	21.355	22.082	23.010	26.206
Lev	43182	0.427	0.211	0.051	0.259	0.416	0.579	0.967
Roe	43182	0.044	0.163	-0.927	0.022	0.063	0.109	0.439
Cfo	43182	0.047	0.069	-0.176	0.009	0.046	0.086	0.255
Growth	43182	0.143	0.409	-0.614	-0.048	0.084	0.238	2.625
Age	43182	11.434	7.975	0.608	4.386	9.707	17.602	29.569
Board	43182	2.111	0.198	1.609	1.946	2.197	2.197	2.639
Indep	43182	37.738	5.403	14.290	33.333	36.363	42.860	57.140
SOE	43182	0.339	0.473	0.000	0.000	0.000	1.000	1.000
Top1	43182	35.123	15.169	7.020	23.490	33.333	45.360	75.460

4.2 Univariate Analysis

Table 3 reports the results of mean difference tests grouped by policy implementation (*Did*=0 and *Did*=1). Regarding the core variable, the mean number of green patent applications (*Patent*) for enterprises in the treatment group was 1.081 following policy implementation, significantly higher than the pre-implementation figure of 0.913, with the difference being significant at the 1% level. This indicates that the implementation of green finance policies led to a significant increase in green innovation output among the treated enterprises. Regarding the control variables, firm size (*Size*), operating cash flow (*Cfo*), growth (*Growth*), board size (*Board*), and the shareholding ratio of the largest shareholder (*Top1*) all showed significant differences before and after the policy implementation, while the differences in the asset-liability ratio (*Lev*) and the nature of property rights (*SOE*) were significant at the 5% and 1% levels, respectively. The above results indicate significant differences across multiple dimensions between the treatment and control groups before and after policy implementation. Therefore, it is necessary to control for these variables in the multiple regression analysis.

Table 3: Univariate Tests

Variable	Before Policy		After Policy		Mean Difference
	Observations	Mean	Observations	Mean	
Patent	41456	0.913	1726	1.081	-0.168***
Size	41456	22.27	1726	22.38	-0.112***
Lev	41456	0.427	1726	0.416	0.012**
Roe	41456	0.0440	1726	0.0480	-0.00400
Cfo	41456	0.0460	1726	0.0570	-0.011***
Growth	41456	0.144	1726	0.106	0.038***
Age	41456	11.43	1726	11.51	-0.0820
Board	41456	2.112	1726	2.089	0.023***
SOE	41456	0.338	1726	0.362	-0.024**
Top1	41456	35.07	1726	36.48	-1.411***

*Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.*

4.3 Regression Analysis

Table 4 reports the baseline regression results. Column (1) controls only for firm and year fixed effects without including additional control variables. The *Did* coefficient is 0.1497 and significant at the 5% level, preliminarily indicating that the policy has a positive impact on corporate green innovation. Column (2) incorporates control variables, such as firm characteristics and corporate governance, into the model used in Column (1). The *Did* coefficient is 0.1366 and remains significant at the 5% level. Although the coefficient decreases slightly, its sign and significance remain consistent, indicating that while the control variables explain a portion of the positive effect, the core conclusion remains unchanged. Column (3) further controls

© 2026 The Author(s). Authors retain copyright in their articles.

for year-industry joint fixed effects to absorb industry-level time-varying shocks. The *Did* coefficient is 0.1143, remaining significant at the 5% level. In terms of control variables, the coefficients of enterprise size (*Size*) and the proportion of independent directors (*Indep*) are both significantly positive, indicating that larger enterprises with better corporate governance structures are more inclined to green innovation, consistent with theoretical expectations.

Table 4: Baseline Regression Results

	Patent		
	(1)	(2)	(3)
Did	0.1497**	0.1366**	0.1143**
	(0.0649)	(0.0615)	(0.0571)
Size		0.3555***	0.3630***
		(0.0184)	(0.0176)
Lev		-0.0466	-0.0305
		(0.0522)	(0.0503)
Roe		0.0280	0.0085
		(0.0286)	(0.0279)
Cfo		0.0111	0.0482
		(0.0665)	(0.0641)
Growth		-0.0009	-0.0088
		(0.0091)	(0.0090)
Age		0.0118	0.0534*
		(0.0269)	(0.0275)
Board		0.0866	0.1094*
		(0.0618)	(0.0574)
Indep		0.0040**	0.0044***
		(0.0018)	(0.0017)
SOE		0.0082	0.0090
		(0.0351)	(0.0339)
Top1		0.0012	0.0013
		(0.0009)	(0.0009)
Cons	0.9138***	-7.4979***	-8.2141***
	(0.0026)	(0.5359)	(0.5191)
Firm FE	YES	YES	YES
Year FE	YES	YES	
Year*Industry			YES
N	43182	43182	43163
adj. R^2	0.704	0.721	0.732

*Note: ***, **, and * indicate that the coefficients are significant at the 1%, 5%, and 10% levels, respectively; t-values are in parentheses.*

4.4 Robustness Tests

4.4.1 Parallel Trends Test

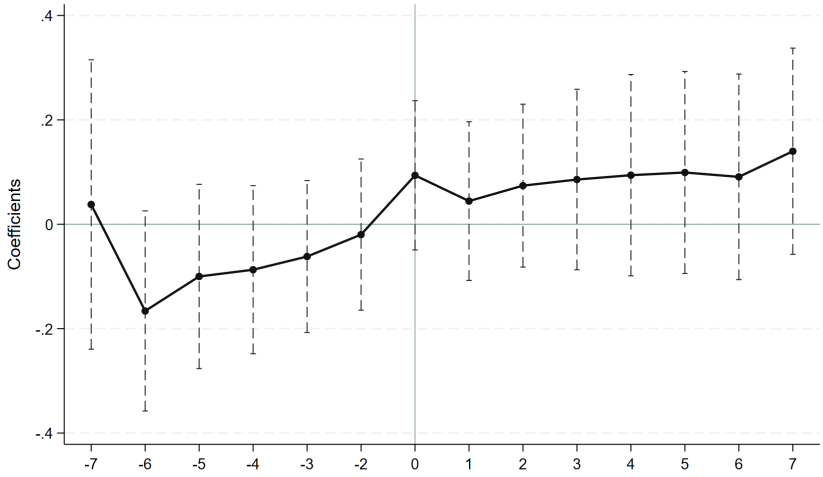
Drawing on the methods of Kong Lingcheng et al. [16], this paper constructs the following dynamic DID model (2):

$$Patent_{i,t}=\alpha+\sum_{k=-7}^{-2}\beta_k D_{i,t}^{k}+\beta_0 D_{i,t}^{0}+\sum_{k=1}^{7}\beta_k D_{i,t}^{k}+\gamma\cdot Controls_{i,t}+\mu_i+\lambda_t+\varepsilon_{i,t} \quad (2)$$

Figure 2 reports the estimated coefficients and 95% confidence intervals for each period. Looking at the periods prior to policy implementation (k = -7 to k = -2), none of the estimated coefficients passed the significance test at the 10% level, and all confidence intervals spanned the zero line. This indicates that prior to the implementation of the green finance pilot zone policy, there was no significant difference in the trends of green patent application volumes between the treatment group and the control group, and the parallel trend test was satisfied.

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 2: Parallel Trends Test



Starting from the period of policy implementation (k=0), the estimated coefficient turned positive and maintained a positive trend throughout the subsequent periods (k=1 to k=7). The policy effect began to intensify gradually from the second period following implementation, and the confidence interval narrowed significantly. This indicates that the policy's role in promoting the green transformation of enterprises is sustained and stable, further validating the findings of the baseline regression analysis.

4.4.2 Placebo Test

In the real-world economic environment, random factors unrelated to policy may interfere with corporate green innovation, leading to biased estimates of policy effects. To rule out this possibility, this paper conducts a placebo test.

Figure 3: Placebo Test

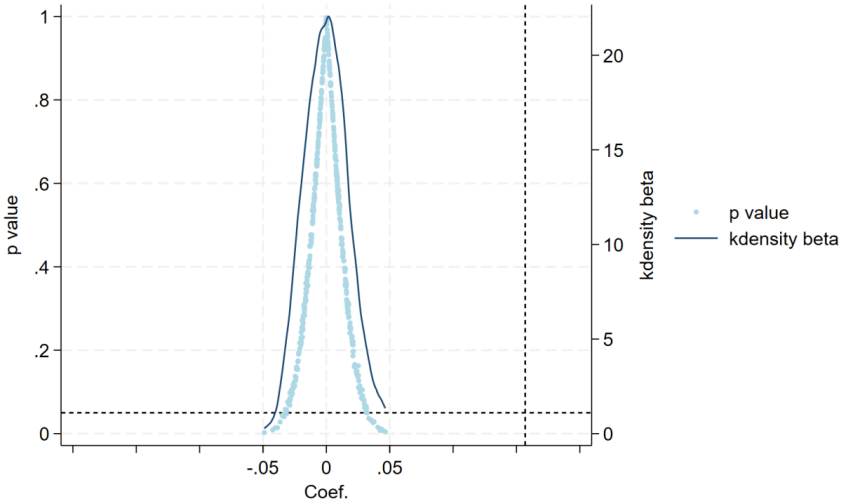


Figure 3 reports the results of the placebo test. An analysis of the coefficient distribution reveals that, following 500 iterations of random matching, the mean of the pseudo-*Did* coefficients approaches 0.000, with a maximum value of only 0.0465; in contrast, the true *Did* coefficient from the baseline regression is 0.1366, a figure far exceeding any of the random estimation results. This means that the true policy effect differs significantly from effects that might arise from random factors. In 500 random estimations, the pseudo-coefficient was significant at the 5% level in only 27 instances, a figure close to the level expected by chance (25 instances). Based on the above results, the impact of unobserved random factors on corporate green innovation is negligible. The promoting effect of green finance policies is not driven by chance, and the placebo test is passed.

© 2026 The Author(s). Authors retain copyright in their articles.

4.4.3 Propensity Score Matching

To mitigate the endogeneity issue arising from sample selection bias, this paper employs the Propensity Score Matching combined with Difference-in-Differences (PSM-DID) method to conduct a robustness test. The variable *Treat* is defined based on whether a firm is affected by the policy: it takes a value of 1 if the firm is registered in the pilot zone, and 0 otherwise. Matching variables include firm size (*Size*), return on equity (*Roe*), operating cash flow (*Cfo*), board size (*Board*), proportion of independent directors (*Indep*), nature of property rights (*SOE*), and the shareholding ratio of the largest shareholder (*Top1*). A Logit model is used to estimate propensity scores, and 1:1 nearest-neighbor matching without replacement is employed, with a caliper set at 0.05.

Before matching, the kernel density curves of the treatment and control groups exhibited certain differences; after matching, the curves largely overlapped, satisfying the common support assumption. Post-matching, the maximum absolute standardized bias for any variable was 3.1%, well below the 20% threshold; t-test p-values exceeded 0.1, indicating no significant differences between the two groups regarding the matching variables and demonstrating effective matching.

4.4.4 Entropy Balancing Matching Test

As a supplementary test, this paper employs the entropy balancing matching method. Entropy balancing weighting is performed using the same matching variables as those used for PSM. Table 5 reports the regression results obtained using PSM-DID and entropy balancing matching. The *Did* coefficient is 0.1312 in Column (1) (PSM-DID) and 0.1257 in Column (2) (entropy balancing matching); both are significant at the 5% level. The results indicate that the core conclusions remain robust after employing various matching methods to mitigate sample selection bias.

Table 5: Robustness Test (PSM-DID and Entropy Balancing Matching)

	PSM Matching	Entropy Balancing Matching
	(1)	(2)
Did	0.1312**	0.1257**
	(2.02)	(2.20)
Controls	YES	YES
Firm FE	YES	YES
Year FE	YES	YES
N	4174	43182
adj. R^2	0.71	0.71

*Note: ***, **, and * indicate that the coefficients are significant at the 1%, 5%, and 10% levels, respectively; t-values are in parentheses.*

4.4.5 4.4.5 Other Robustness Tests

Table 6: Other Robustness Tests

	Differentiating Policy Timing (First)	Differentiating Policy Timing (Second)	Differentiating Policy Timing (Third)	Replacing the Dependent Variable
	(1)	(2)	(3)	(4)
Did	0.1542**	0.1345	0.0783	0.0516**
	(2.01)	(0.72)	(0.66)	(2.09)
Controls	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
N	42383	40571	40932	43182
adj. R2	0.72	0.72	0.72	0.49

*Note: ***, **, and * indicate that the coefficients are significant at the 1%, 5%, and 10% levels, respectively; t-values are in parentheses.*

(1) Distinguishing Policy Timing

Given that the pilot zones were established and rolled out in batches, pooling all batches for regression analysis might mask the heterogeneity of the policy effects. To this end, this paper excludes samples from the other batches and re-estimates the DID model by retaining only the pilot enterprises from a single batch and the enterprises that were never affected by the policy. Columns (1) through (3) of Table 6 show that the *Did*

© 2026 The Author(s). Authors retain copyright in their articles.

coefficient for the first batch is 0.1542 (significant at the 5% level), while the coefficients for the subsequent two batches are 0.1345 and 0.0783, respectively; neither of the latter two reaches the level of statistical significance. This indicates that the first batch of policies yielded the most significant results, while the effects of the subsequent two batches were relatively weaker, a pattern that may be attributable to policy transmission lags or diminishing marginal effects.

(2) Replacing the Explained Variable

To avoid potential bias in the estimation results arising from the measurement method of the dependent variable, this paper converts the continuous patent count into a dummy variable (*patentdummy*), indicating whether a firm held a patent in a given year, to examine the marginal impact of the policy. Column (4) of Table 6 shows a *Did* coefficient of 0.0516, significant at the 5% level, indicating that the policy significantly increased the probability of firms engaging in innovation activities. This is consistent with the baseline regression results.

5. Heterogeneity Analysis

5.1 Governance Effect: Management Shareholding Ratio

Different levels of internal corporate governance will have different responses to policy effects. Existing research has shown that management ownership concentration will inhibit the green innovation of enterprises. Management shareholding can alleviate agency conflicts between large shareholders and small and medium shareholders, thereby promoting enterprises to carry out more green technology innovations [17]. Referring to its approach, this article puts the management shareholding ratio (*Mngmhldn*) and its cross-product term *Did* × *Mngmhldn* with Did into model (1) for regression. The results are shown in column (1) of Table 7.

It can be seen from column (1) of Table 7 that the coefficient of the cross-multiplication term of *Did* × *Mngmhldn* is significantly negative at the 10% level (estimated coefficient = -0.005, t value = -1.86). The green finance reform and innovation experimental zone policy will have a weaker promotion effect on green innovation in companies with a high management shareholding ratio. This may be because the green finance reform and innovation pilot zone policy forces companies to standardize environmental behavior, increase investment in green innovation, and improve corporate governance through market-oriented mechanisms such as green credit guidance and environmental protection regulatory constraints. For companies with low management shareholding ratios and weak internal governance, the effect of policies on improving governance levels is more prominent, which is in line with the expectations of this article.

5.2 Equity Effect: Government Green Subsidies

Under the traditional policy framework, certain enterprises, due to factors such as scale, ownership structure, or ties between government and business, find it easier to secure government subsidies, creating implicit barriers for small and medium-sized enterprises. Yu Zhimai found that government environmental subsidies exert a significant negative moderating effect on policy effectiveness; specifically, the lower the subsidy level, the stronger the policy's role in promoting corporate green innovation [10]. Drawing on the research approach of those studies, this paper introduces a variable, government green subsidy (*Green_Subsidy*), defined as the amount of government subsidies related to environmental protection received by the firm during the year. Government green subsidies (*Green_Subsidy*) and their interaction term with DID (*DID* × *Green_Subsidy*) were included in Model (1) for regression analysis. The results are presented in Column (2) of Table 7.

Column (2) of Table 7 reports that the interaction term *Did* × *Green_Subsidy* is significant at the 10% level (estimated coefficient = -0.096, t-value = -1.71). The results indicate that for enterprises that had already received substantial green subsidies, the policy's role in promoting their green innovation was actually weaker. It can therefore be inferred that the policy has, to some extent, enhanced market fairness and lowered the implicit barriers for enterprises to access green financial resources. By accommodating marginalized entities, such as small and medium-sized enterprises (SMEs) and non-state-owned enterprises, it has provided them with more equitable opportunities for green financing.

© 2026 The Author(s). Authors retain copyright in their articles.

5.3 The Amplification Effect of Media Attention: Media Attention Level

As important information intermediaries in the capital market, the media play a corporate governance role by influencing the reputations of managers [18]. Media attention plays a significant supervisory and governance role across various aspects of the behavior of listed companies in China [19]. Drawing on the approaches of Li Peigong and Shen Yifeng, Xu Liping and Li Yu, and Huang Jun and Guo Zhaorui [20–22], this study measures media attention (*Media*) as the natural logarithm of (1 plus) the number of headlines featuring the firm in newspapers and online media during the year. We incorporate *Media* and the interaction term *Did*×*Media* into Model (1) for regression analysis The results are presented in Column (3) of Table 7.

The results show that the *Did*×*Media* interaction term is significant at the 5% level (estimated coefficient = 0.184, t-value = 1.99). This indicates that media coverage amplifies the impact of policies by mitigating information asymmetry between firms and financial institutions, as well as by strengthening external oversight and reputational constraints. The higher the level of media attention, the more fully policy information is disseminated, the more actively enterprises respond to policies, and the greater the output of green innovation.

Table 7: Heterogeneity Analysis

	(1)	(2)	(3)
	Patent	*Patent*	*Patent*
Did	0.156**	0.157**	0.112*
	(2.20)	(2.53)	(1.65)
Mngmhldn	0.002**		
	(2.36)		
Did×*Mngmhldn*	-0.005*		
	(-1.86)		
Green_Subsidy		0.029**	
		(2.41)	
Did×*Green_Subsidy*		-0.096*	
		(-1.71)	
Media			0.021
			(1.43)
Did×*Media*			0.184**
			(1.99)
Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	43021	43182	42602
adj.R²	0.7209	0.7212	0.7210

6. Conclusions and Policy Recommendations

Since 2017, the People's Bank of China and six other government departments have established pilot zones for green finance reform and innovation across selected provinces and regions in three batches, using market-based mechanisms to channel financial resources toward green innovation. Using a sample of non-financial A-share listed companies on the Shanghai and Shenzhen stock exchanges from 2011 to 2024, this paper employs a multi-period difference-in-differences model to empirically examine the impact of pilot zone policies on corporate green innovation. Research finds that the policies in the pilot zones have significantly boosted the green innovation levels of enterprises within their jurisdictions, as evidenced by an increase in the number of green patents. Further research reveals that enterprises with low management shareholding ratios are more significantly affected by the policy, indicating that the policy can compensate for governance deficiencies. Enterprises that received substantial green subsidies beforehand were less affected by the policy, demonstrating the policy's equitable effect. Enterprises attracting greater media attention are more strongly affected by policies, confirming that the media may exert an amplifying effect by mitigating information asymmetry and reinforcing reputational constraints.

Based on the above conclusions, this article makes the following suggestions. From a policy perspective, first, the innovation of green financial products and services should be continuously deepened. Pilot zones should further expand the scope of financing secured by environmental rights, such as carbon emission allowances and pollution

© 2026 The Author(s). Authors retain copyright in their articles.

discharge rights, lower financing thresholds for green R&D by small and medium-sized enterprises (SMEs), and more fully realize the equitable effects of these policies. Second, improve environmental information disclosure and media supervision mechanisms. Regulatory authorities should accelerate the establishment of a mandatory disclosure system for corporate environmental information, use media reports to reduce information asymmetry between companies and financial institutions, and strengthen reputation constraints and social supervision, thereby amplifying the market transmission effect of policies. Third, establish a mechanism for the dynamic assessment of policy effects and for feedback. Given that the impact was most significant in the first batch of pilot zones, with diminishing returns observed in subsequent batches, regulatory authorities should periodically evaluate the effectiveness of policies and the industrial structures within each zone, as well as the actual needs for green transformation, in order to dynamically adjust the policy mix and intensity.

At the corporate level, the first step is to optimize management incentives and internal governance structures. Enterprises should proactively refine mechanisms for management shareholding and incentive-constraint systems, foster management recognition of the long-term strategic value of green innovation, and leverage the institutional dividends and financing conveniences offered by policies. Secondly, enterprises should proactively tap into market-based financing channels. Corporate management should proactively enhance the quality of environmental information disclosure and capabilities in green technology R&D, leveraging strong green performance to secure greater credit support and actively drive the green transition. Finally, enterprises should formulate differentiated strategic plans for green innovation. They should objectively assess their own conditions, such as governance standards, resource endowments, and the information environment, to select targeted green innovation pathways and effectively translate policy dividends into sustainable green competitive advantages.

References

[1] Si, L. J., & Yao, X. Q. (2022). Green finance reform and innovation and the ecologicalization of regional industrial structure: A quasi-natural experiment from green finance reform and innovation pilot zones. Study & Exploration, (4), 129–138+2.

[2] Mao, X. M., & Wang, R. Z. (2024). Green finance and new quality productive forces: Promotion or inhibition? From the perspectives of technological innovation and environmental attention. Journal of Shanghai University of Finance and Economics, 26(5), 30–45. https://doi.org/10.16538/j.cnki.jsufe.2024.05.003.

[3] Hu, J. B., Ye, S., & Zhang, K. Y. (2024). Green finance policy and corporate ESG performance: A quasi-natural experiment based on green finance reform and innovation pilot zones. Finance Forum, 29(7), 59–69. https://doi.org/10.16529/j.cnki.11-4613/f.2024.07.005.

[4] Wang, Y., Wang, J. Z., Wu, Z. S., & Wang, W. W. (2025). Can green finance promote the economy's shift from virtual to real? Empirical evidence based on green finance reform and innovation pilot zones. Journal of Financial Research, (12), 20–38.

[5] Wang, X., & Wang, Y. (2021). Research on green credit policy promoting green innovation. Journal of Management World, 37(6), 173–188+11. https://doi.org/10.19744/j.cnki.11-1235/f.2021.0085.

[6] Li, W. J., & Zheng, M. N. (2016). Substantive innovation or strategic innovation? The impact of macro-industrial policies on micro-enterprise innovation. Economic Research Journal, 51(4), 60–73.

[7] Lu, G. Q., Wang, Z., & Zhang, C. Y. (2014). Research on the performance of government innovation subsidies for China's strategic emerging industries. Economic Research Journal, 49(7), 44–55.

[8] Dong, Z. Y. Z., Wei, S. J., & Zhang, W. W. (2025). The green innovation effect of energy conservation and emission reduction. China Economic Quarterly, 25(3), 767–784. https://doi.org/10.13821/j.cnki.ceq.2025.03.14.

[9] Liu, J. K., & Xiao, Y. Y. (2022). China's environmental protection tax and green innovation: Leverage effect or crowding-out effect? Economic Research Journal, 57(1), 72–88.

© 2026 The Author(s). Authors retain copyright in their articles.

[10] Yu, Z. M. (2021). Environmental interviews, government environmental subsidies and corporate green innovation. Foreign Economics & Management, 43(7), 22–37. https://doi.org/10.16538/j.cnki.fem.20210508.201.

[11] Yu, X. G. (2025). The green innovation effect of inter-enterprise environmental protection behavior: A study based on the national private enterprise survey. Guizhou Social Sciences, (4), 126–134. https://doi.org/10.13713/j.cnki.cssci.2025.04.015.

[12] Lu, J., Yan, Y., & Wang, T. X. (2021). Research on the micro effects of green credit policy: From the perspectives of technological innovation and resource reallocation. China Industrial Economics, (1), 174–192. https://doi.org/10.19581/j.cnki.ciejournal.2021.01.010.

[13] Li, S. H., You, B. Y., Niu, K., & Yin, J. W. (2025). Green credit, informal financial financing and green innovation of heavily polluting enterprises. Economic Research Journal, 60(9), 198–217.

[14] Guo, J. J., Fang, Y., & Guo, Y. (2024). Environmental regulation, short-term failure tolerance and corporate green innovation: Evidence from the practice of green credit policy. Economic Research Journal, 59(3), 112–129.

[15] Wang, S. Q., & Wang, Y. Q. (2025). Research on the impact of green credit policy on green innovation of heavily polluting enterprises. Journal of Industrial Technological Economics, 44(3), 67–77.

[16] Kong, L. C., & Chai, Z. Y. (2021). Does the upgrading of provincial development zones improve urban economic efficiency? Quasi-experimental evidence from heterogeneous development zones. Journal of Management World, 37(1), 60–75+5. https://doi.org/10.19744/j.cnki.11-1235/f.2021.0005.

[17] Wu, S. C., Qu, D., Guo, Y., & Dong, J. C. (2023). Ownership concentration, management shareholding and corporate green technology innovation. Finance Research, (6), 80–89. https://doi.org/10.14115/j.cnki.10-1242/f.2023.06.009.

[18] Dyck, A., Volchkova, N., & Zingales, L. (2008). The corporate governance role of the media: Evidence from Russia. The Journal of Finance, 63(3), 1093–1135. https://doi.org/10.1111/J.1540-6261.2008.01353.X.

[19] Kong, D. M., Liu, S. S., & Ying, Q. W. (2013). The role of the media in corporate behavior: Clearing the muddy waters or adding fuel to the fire? Journal of Management World, (7), 145–162. https://doi.org/10.19744/j.cnki.11-1235/f.2013.07.013.

[20] Li, P. G., & Shen, Y. F. (2010). The corporate governance role of the media: Empirical evidence from China. Economic Research Journal, 45(4), 14–27.

[21] Xu, L. P., & Xin, Y. (2011). Media governance and protection of minority investors. Nankai Business Review, 14(6), 36–47+94.

[22] Huang, J., & Guo, Z. R. (2014). News media coverage and capital market pricing efficiency: An analysis based on stock price synchronicity. Journal of Management World, (5), 121–130. https://doi.org/10.19744/j.cnki.11-1235/f.2014.05.010.

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.

© 2026 The Author(s). Authors retain copyright in their articles.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.3346
Published by: Zeus Press

Research on the Impact of High-Speed Rail Opening on Financial Resource Allocation in the Context of Urban-Rural Integration: An Estimation Based on the PSM-DID Model

Jiaying Li*

School of Finance, Shanxi University of Finance and Economics, Taiyuan 030006, China

**Corresponding author: Jiaying Li.*

Abstract

Over the past few years, due to the expansion of the high-speed rail network in China and the ongoing development of rural revitalization and urban-rural integration, the flow of people, money, and other factors has been increased by reducing the distances over space and time between cities, which deeply changes the structure of regional financial resource allocation. This paper uses the PSM-DID model to empirically examine the effect of high-speed rail opening on financial resource allocation in prefecture-level cities and its transmission mechanism using panel data of 300 prefecture-level cities across the country in 2007-2022. The baseline regression results show that high-speed rail opening has a strong and positive promoting role on regional financial development, and this finding holds true even after conducting a number of robustness analyses, including placebo tests. Heterogeneity analysis shows that the promoting effect of high-speed rail on financial resource allocation is stronger in the east region, large and medium-sized cities, and ordinary prefecture-level cities. In terms of the optimal configuration of the high-speed railway network, enhancing the absorption capacity of underdeveloped areas, adopting differentiated policies, balancing regional finance resource allocation, and developing supporting guarantee systems, this paper offers recommendations to solve the financial siphon effect brought about by the high-speed railway and aims to offer references regarding how to promote the balance of urban and rural financial development and facilitate the construction of high-speed railways to support rural revitalization and urban-rural integration.

Keywords

high-speed rail opening, financial resource allocation, urban-rural integration, rural revitalization

1. Introduction

The revitalization of rural areas and the integrated development of urban and rural areas are the central policies of the state in coordinating the development of regions and promoting the growth of common prosperity in China. Nevertheless, the persistent imbalance of urban-rural dual structures is still significant. According to the information provided by the Ministry of Agriculture and Rural Affairs, there is still a supply gap in rural financial services at the county level, where rural areas have inadequate capital supply, poor coverage of offline financial services, and it is expensive to obtain finance for specialty agriculture and rural

© 2026 The Author(s). Authors retain copyright in their articles.

industries. Being the key factors of production of the modern economy, financial resource spatial allocation balance and efficiency directly affect the pace of rural industrial upgrading and the magnitude of both-ways factor flows between urban and rural zones; thus, act as the main lever of closing the urban-rural development gap and the implementation of urban-rural integration policies.

The transport infrastructure plays an essential role in reshaping the spatial structure of regional elements and controlling the trend of financial resources streams. Economic spillovers of high-speed rail (HSR), as one of the most powerful drivers of modern transport infrastructure development in China, have been the subject of much research by both policymakers and academics [1, 2]. The No. 1 Central Document of 2025 explicitly calls for promoting rural revitalization and improving rural infrastructure through reform, opening-up, and technological innovation [18]. At present, China's HSR network construction has entered a new stage of high density and full coverage. By the end of 2024, the national HSR operating mileage exceeded 48,000 km, covering cities with a population of over 500,000 in 97% of the country's regions [19], and the "Eight Vertical and Eight Horizontal national HSR backbone network" has been built. The effect of time-space compression and factor-acceleration introduced by HSR does not only impact the macroeconomic level of urban agglomerations, but also penetrates into the meso-and-micro levels of urban-rural and county-level areas, which become significant factors of regional economic space structure change, and contribute to the rearrangement of population, capital, and industrial factors between urban and rural areas [3, 4].

The opening of HSR has two negative but different logical consequences in the allocation of urban-rural financial resources, which are positive power, and siphoning/bleeding impacts. On the one hand, HSR eliminates the urban-rural transportation barriers, minimizes the transportation and informational expenses of cross-regional operations, and the extension of the offline branches of the financial institution, encourages commercial banks, rural banks, and other financial institutions to open offices in counties and towns, diversifies the rural deposit and loan business channel, scales up the size of rural financial services supply, and rationalizes the balanced urban-rural financial resource allocation pattern. Alternatively, HSR significantly reduces the cost of inter-city commuting and capital flow, which enhances the advantages of central and medium-large cities in terms of factor agglomeration, which leads to a strong financial siphoning effect. In particular, the high-quality labour and industrial capital of the countryside continues to migrate to large cities, and credit quotas and branch resources of banks are shifted towards the centers of the city, which further increases the outflow of financial resources in rural areas and creates an even larger urban-rural financial development gap.

The review of available related literature shows that there are still significant gaps and limitations to the existing research. Firstly, most of the current literature is based on the effects of opening the HSR on the overall macroeconomic variables, including regional GDP growth, industrial agglomeration, and per capita household income, whereas systematic studies that concentrate on the changes in the spatial distribution of financial resources between city and countryside and their financial supply arrangements during the shock of HSR are less common. Second, the majority of the studies on HSR-finance are performed on individual cities or urban agglomerations, with the research focus limited to intra-urban financial development; comparative analyses that compare urban and rural aspects using the urban-rural dual structure are absent, and therefore the various consequences of the HSR on urban-rural finance resource reallocation cannot be comprehensively determined.

According to the policy context provided above, existing contradictions and gaps in research, this paper chooses the panel data of 300 prefecture-level cities throughout China between 2007 and 2022, uses the opening of HSR as a quasi-natural experiment, and uses the PSM-DID model to deal with the problem of sample selection bias and endogeneity. This methodically determines the net impact of HSR opening on the magnitude of regional financial resource allocation, elucidates three main transmission channels, i.e., population concentration, industrial development, and urbanisation synergy, and performs heterogeneity tests by region, city size, and administrative hierarchy to numerically determine various characteristics of HSR financial empowerment effects. To meet the challenge of rural financial resource loss due to HSR siphoning, different policy optimisation strategies are suggested. The marginal contributions of this paper compared to the current literature are threefold: (i) it has deviated from the conventional single-regional macro-analytical approach and examines how HSR affects the urban-rural dimensional financial resource allocation based on the dual strategic views of the country, i.e. urban-rural integration and rural revitalization, which fills the research gap in the urban-rural binary perspective; (ii) in terms of research content, the research considers both the positive effect of agglomeration and the negative effect of financial siphoning simultaneously, and

© 2026 The Author(s). Authors retain copyright in their articles.

identifies the dual logic of action and multi-dimensional heterogeneity characteristics across eastern/central-western areas, large/small-medium cities, and major cities/provincial capitals/ordinary prefecture-level cities; (iii) in terms of research mechanisms, it thoroughly examines the transmission mechanisms through three pathways, including population, industry, and urbanization, and adds fiscal decentralization as a moderating constrain variable, which enriches the empirical analytic paradigm in the interdisciplinary field of transport and finance.

The rest of the paper will be organized in the following way: Chapter 2 contains the literature review, Chapter 3 contains the theoretical analysis and the mechanisms of influence, Chapter 4 outlines the empirical design of the research, Chapter 5 provides the empirical analysis, and Chapter 6 provides conclusions and policy recommendations.

2. Literature Review

The existing academic study of the consequences of the high-speed rail opening on the financing of resources is quite profound and may be roughly divided into three groups: the first one focuses on the consequences of high-speed rail opening on regional development imbalance; the second one studies the changes in financial resource flows against the backdrop of high-speed rail opening; and the third one considers the process mechanisms of how the siphon effect of high-speed rail affects financial resource allocation.

2.1 Research on the Impact of High-Speed Rail Construction on Regional Economic Development

The scholarly society today has developed quite an extensive literature on the economic implications of building high-speed railways with the researches largely concentrating on the macroeconomic effects of the construction of high-speed railways, coordinated growth of urban agglomerations as well as the changes in regional disequilibrium [2, 4].

Based on the coordinated development of urban agglomerations, the authors (Lan Xiujuan and Zhang Weiguo, 2023) used the nine largest urban agglomerations in China as their empirical examples, and they examined with great detail the inner mechanisms of the influence of high-speed rail networks on the economic coordination of urban agglomerations. Their results showed that the reinforcement of high-speed rail networks increases the internal economic connections between urban agglomerations, where the networked agglomeration structure focused on core cities is growing outwards [3]. In the course of such a process, the agglomeration capacity of the majority of cities was increased; nevertheless, the obstacles to economic linkages across provincial administrative borders were not eliminated thoroughly, and inter-regional cooperation was still rather weak. Even more so, considering network density dynamics, it was found that high-speed rail networks enhanced agglomeration effects, yet increased disparities in intra-urban-agglomeration economic development, which had some restraining effect on the coordinated and sustainable growth of regional economies.

The Yangtze River Delta urban agglomeration has been studied by Yan Yalei et al. (2023), and they have performed quantitative analysis in two ways, which are regional accessibility and spatial linkage patterns [4]. It was indicated in the study that the construction of the high-speed rail network changed the railway organization of the Yangtze River Delta from an inverted Z-shaped form to a flat, hub-centered radial model. The reconstruction created new optimal allocation routes of regional factors, rearranged the system of territorial division of labour, and gave a boost to the regional development of high quality. Nevertheless, such reorganization was not entirely positive because it triggered the potential for development of core regions, but, at the same time, objectively increased the disparity between development levels of different regions.

Wang Yufei and Ni Pengfei (2016) were some of the first people to suggest the so-called siphoning effect due to high-speed rail [2]. The authors verified that high-speed rail could significantly reduce time and space distances between cities, increase the spatial spillover impacts of the economic activities of an urban area, and prove the role of transport infrastructure construction in economic development. However, this promotional impact was spatially asymmetric. In macro-layout at a national level, it meant that the eastern and central regions with stronger economic backgrounds were hence brought within the sphere of core development, whereas the northeastern and western regions with poorer economic backgrounds were endangered by marginalisation.

© 2026 The Author(s). Authors retain copyright in their articles.

Li Hongchang (2019) also analyzed the effect of the siphon effect on small and medium-sized cities. He brought in the main idea of economic gradient effects and stated that major spatial gradient differences are evident in urban agglomeration economies in China. The high-speed railway did not slow but hastened the realization of gradient effects, which saw factor resources flow in one direction to the central cities, thus, restricting the creation and growth of industries in small and medium-sized cities and simultaneously resulting in the disappearance and loss of critical factors of production like talents [5].

According to the available literature on high-speed rail opening, it is possible to state that there has been developed a multifaceted analytical structure of the economic impact of high-speed rail. This framework not only recognizes the positive importance of high-speed rail in the reorganization of spatial patterns and facilitation of factor flows but also carefully notes its negative contribution to increasing the gap between regions and supporting central agglomeration, therefore, it offers a good theoretical background to the follow-up investigation.

2.2 Research on Financial Resource Allocation

The scholarly study of financial resource allocation can be described as in-depth studies based on conceptual definitions and the most important factors that affect it. In relation to the interrelationships between financial structure, regional industrial development, and economic growth, Fan Fangzhi and Zhang Lijun (2003) began with the internal connection of financial structure change and industrial structure change, and utilized panel data of eastern, central, and western China over the period 1978-2000 to perform empirical analysis of the dynamism of the three [6]. The results of their study showed that the lower degree of development of financial structure and the slower rate of structural change in the central and western regions of China contributed significantly to the inability to optimize and upgrade industrial structures and achieve sustained economic growth within those regions. On the contrary, the delay of the industrial structure upgrading process in the central and western regions also hindered the gradual progress of the regional financial industry, which, in turn, worsened the underdevelopment of local economies and created a self-reinforcing cycle constraint system.

From the viewpoint of regional economic growth, Zhou Li (2002) claimed that the effectiveness of financial resources allocation is a decisive factor in regional economic development. That is to say, he indicated that development of the regional financial sector could bring about a continuous impulse to support the long-term economic growth, and that disparities in the level of financial development among regions also play a role in the divergence of regional economic growth in China [7]. It was also observed in this research that when regional financial development commences late and under unfavorable conditions, its influence on long-term economic development is likely to be more severe. Of all the financial development measures, the most important was the positive correlation between the degree of financial marketization and economic growth. Zhao Jian (2025), based on the integration of digital finance and the real economy, explored the mechanisms of how digital finance helps to facilitate high-quality development in the manufacturing sector [8]. The empirical findings revealed that digital finance has a positive influence on high-quality manufacturing development as it helps to quickly agglomerate the factors of innovation to reduce the misallocation of factors of labour and capital, which leads to the improvement of the efficiency of total factor resource allocation and eventually improves the general level of high-quality manufacturing development.

2.3 The Impact of High-Speed Rail Opening on Financial Resource Allocation

In terms of the influence of opening a high-speed railway on the allocation of financial resources, the available literature has taken a look at this issue in many ways and has used diverse empirical studies, with the first one being the relationship between opening a high-speed railway and financial resource allocation, and subsequent in-depth explanations of the mechanisms and heterogeneity of both the effects of opening a high-speed railway on financial agglomeration, corporate financing costs, online inclusive finance growth, and so on.

The Cheng Jiawen and Cui Huihui (2025) used a continuous-time difference-in-differences approach to empirical testing, which mainly focused on the influence of high-speed rail opening on the perception of housing prices among the population and explored the intermediate transmission mechanism of financial agglomeration [9]. The findings revealed that the opening of high-speed rail could be effective in improving

© 2026 The Author(s). Authors retain copyright in their articles.

the regional locational competitiveness, which leads to the attraction of different financial resources that would agglomerate within the region, which eventually influences the real estate market through the expansion of credit scale and cross-regional capital flows. It also indirectly reflected the change trend of regional financial resource allocation after the opening of high-speed rail. Wu Ying et al. (2021) applied a difference-in-differences model to investigate how the cost of debt financing changed in micro-enterprises following the introduction of high-speed rail [10]. The researchers concluded that the opening of high-speed rail can lead to substantial savings in corporate debt financing costs, and it is achieved by increasing the intensity of banking competition in areas where high-speed rail can be accessed. Also, the reduction effect was even stronger among enterprises with less strong financing restrictions and in regions with lesser degrees of financial marketization. Li Xin and Li Fuping (2021) used a multi-period difference-in-differences approach with prefecture-level cities in China as the objects of research to empirically determine the influence of high-speed rail opening on the development of regional digital inclusive finance [11]. Their work showed that the opening of high-speed rail has a highly positive effect on the level of development of regional digital inclusive finance, and this result also held true when methods were applied to address the problem of self-selection. The research also divided prefecture-level cities into third-tier, fourth-tier, and fifth-tier cities based on the Yicai city classification criteria and performed a heterogeneity analysis. The findings indicated that the promoting effect of high-speed rail opening existed solely in third-tier cities, with no statistically significant effect in fourth-tier or fifth-tier cities. Besides, among the fourth-tier and fifth-tier cities (both considered to be lower-tier cities), there were differences in the promotional effects of high-speed rail opening on the development of digital inclusive finance between the two tiers.

3. Theoretical Analysis and Impact Mechanisms

The reasoning behind the consequences of the opening of high-speed rail (HSR) on financial resource allocation is based on the time-space compression effect that destroys both geographic and informational barriers and enables free movement of production factors across regions, including population, capital, and industry, and changes the nature of regional factor agglomeration and diffusion. It causes shifts in financial services supply to demand, enhances the structural underpinning of financial growth to supply, and eventually leads to the spatial redistribution of financial resources. Nevertheless, the financial empowerment effect of HSR opening is not equal in all regions since it is limited to regional foundational conditions, institutional environment, and other factors and therefore results in different financial resource allocation outcomes.

In accordance with the concept of growth poles, economic development tends to be initiated in a well-developed polar region and subsequently spread and affect other regions via spread and backwash effects [12]. A transport network of a modern type, such as HSR, has the potential not only to support the growth pole role of central cities but also to contribute to the growth of peripheral areas by virtue of its spread effect [13]. HSR reduces inter-city time distance significantly and minimizes the costs of time and economy associated with population movement between regional boundaries, and hence, leads to greater regional population sizes and densities, resulting in rigid growth in the demand for financial services, including household consumption, entrepreneurship, and investment financing. The market demand will force financial institutions to extend their offline networks, constantly add more bank branches, and directly increase the volume of transactions in the regional deposit and loan markets, which will lead to the creation of an endogenous momentum of supply-demand-induced supply in the development of finance and finally shift the allocation of financial resources to population-dense regions.

Krugman in the 1990s offered new economic geography, which suggests that there is a natural interaction between transport costs, economies of scale, and spatial agglomeration. The HSR significantly decreases the cost of regional transportation and the cost of information communication, changing industrial location and capital movement routes [14, 15]. The opening of HSR improves the regional economy and increases economic density, leading to the upgrading of regional industrial structures. On the one hand, the upgrading of the regional industries offers the financial institutions many high-quality credit customers and the financial carriers, thus enhancing banks' readiness to lend money and their ability to take deposits, and making the process of the allocation of financial resources more efficient. On the other hand, the cross-region flows of factors brought about by industrial upgrading constantly trigger the vitality of financial market transactions, enhance the growth of the financial resource scale and cross-industry and cross-region optimum distribution, and offer sound real-economic resources to the development of regional finance.

© 2026 The Author(s). Authors retain copyright in their articles.

Based on the core-periphery theory, the natural traits of regional development are spatial polarisation, in which the core cities tend to accumulate resources, and the peripheral regions tend to witness outflows of factors. The enhancement of HSR systems increases the resource-control capability of core cities at the same time as changing the urban-rural spatial linkages [16]. Financial activities are intrinsically spatially differentiated, with financial resource flows being limited by geographical distance and information asymmetry, and this is enhanced by the viewpoint of financial geography [17]. Information transmission is made faster with the help of HSR, geographical barriers are weakened, and urban-rural financial linkage channels are opened. The opening of the HSR not only stimulates the expansion of urban space but also functional upgrading, and constantly increases the urbanization rate of the region. Urbanisation will enhance the urban infrastructure, the financial literacy of the residents, and increase the scope of financial services and the financial product demand. A certain existing level of regional financial development offers a developed market environment to the empowerment effect of HSR implementation, minimizes the market losses of financial institution structure and credit distribution, enhances the driving effect of HSR on bank branch formation and deposit-loan scale growth, and makes it easier to allocate financial resources between urban and rural areas and between cities.

Hypothesis 1: HSR exerts a positive impact on financial resource allocation.

Hypothesis 2: HSR positively affects financial resource allocation through the population factor agglomeration mechanism.

Hypothesis 3: HSR positively affects financial resource allocation through the industrial upgrading mechanism.

Hypothesis 4: HSR positively affects financial resource allocation through the urbanisation mechanism.

4. Empirical Research Design

4.1 Variable Description and Model Specification

4.1.1 Variable Description

The main explanatory variable is the high-speed rail (HSR) opening status, which is a dummy variable equal to 1 in the year when a city opens an HSR line and later and 0 otherwise. The interaction term DID (if the opening year is early) is a dummy variable that equals early opening years.

The independent variables were chosen as the sum number of all the bank outlets, the logarithm of the year-end balance of loans extended by the financial institutions, and the logarithm of the year-end balance of deposits made by the financial institutions, which act as proxy indicators of regional financial development, measuring physical coverage of financial services, the amount of capital supplied, and the amount of capital agglomerated, respectively.

Control variables that are used in the study include population size, population density, economic density, economic development level, financial development level, openness to the outside world, the rate of urbanisation, the level of fiscal decentralisation, industrial structure upgrading, human capital level, and other factors, such as commodity retail sales, which can be used to control the effect of regional development conditions related to the foundations of financial development on financial development.

4.1.2 Sample Selection and Data Sources

The selected data used in this research are the financial development level data on 300 prefectural cities of the whole country based on the information contained in the China City Statistical Yearbook, which covers the years between 2007 and 2022 (the missing raw data is filled with linear interpolation as far as possible). The number of cities is about 300. HSR opening data comes from publicly accessible data of China State Railway Group Co., Ltd.

4.1.3 Model Specification

The use of the Difference-in-Differences (DID) method with the Propensity Score Matching (PSM-DID) approach is used to reduce the problem of sample selection bias and endogeneity as well as to examine the

© 2026 The Author(s). Authors retain copyright in their articles.

effect of X on Y. In the model specification, city fixed effects and year fixed effects are controlled, and statistical inference is done with the help of standard errors clustered at the city level. The exact model specification is presented in the table underneath.

$$Finance_{it} = \beta_0 + \beta_1 DID_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

Where the financial factor flow indicator (total number of bank outlets / ln(loan balance) / ln(deposit balance)) for city *i* in year *t*; DID_{it} is the core explanatory variable, representing the interaction term for HSR opening; $Controls_{it}$ is the set of city-level control variables; μ_i denotes individual fixed effects, controlling for city heterogeneity; λ_t denotes year fixed effects, controlling for time trends; and ε_{it} is the random disturbance term, with standard errors clustered at the city level.

4.2 Descriptive Statistical Analysis

On the basis of the results of descriptive statistics of the numerical variables of 4,298 observations in Table 1, it can be stated that the distribution of the main variables and the regional conditions of foundation of development are given as follows.

Table 1: Descriptive Statistics

Unnamed: 0	Observations	Mean	Median	Std. Dev.	Min	Max
Years since high-speed rail opening	4298	2.9255	1	3.812	0	14
Total number of bank branches	4298	406.957	298	495.4867	3	3376
Population size	4298	6.0348	5.9797	0.6608	4.2536	7.7075
Population density	4298	5.8851	5.9572	0.9095	2.2464	7.7504
Economic density	4298	7.0836	7.2824	1.5094	0.5679	9.8535
Economic development level	4298	2.6906	0.7256	3.7494	0.2286	9.9777
Financial development level	4298	2.0256	2.2657	1.5512	0.2385	6.8068
Openness to foreign trade	4298	0.5194	0.4697	0.2252	0.2244	0.9908
Urbanization rate	4298	0.4828	0.434	0.2149	0.2241	0.9874
Fiscal decentralization	4298	0.5179	0.4793	0.2065	0.2254	0.9798
Industrial structure upgrading	4298	0.7515	0.6837	0.5086	0.2344	3.5411
Human capital level	4298	0.5453	0.5248	0.2294	0.2249	0.9888
Retail sales of commodities	4298	0.4087	0.3824	0.1247	0.2273	0.9349
ln(Year-end financial institution loan balance)	4317	16.3638	16.2214	1.2615	13.8867	19.8065
ln(Year-end financial institution deposit balance)	4317	16.7788	16.6799	1.1305	14.4991	20.0564

5. Empirical Analysis

5.1 Analysis of Baseline Regression Results

5.1.1 DID Baseline Regression Results

Columns (1) to (3) in Table 2 show the subsequent addition of the control variables. The core explanatory variable DID coefficient is significantly positive at the 1 percent level all along; when all the variables are included, the coefficient is 9.36, which means that the opening of the high-speed rail has a significant positive impact on the financial development of a region (number of bank branches used as a proxy variable).

The size of the population, the level of economic development, the level of financial development, the pace of urbanization, and other variables are all important positive determining factors, and they form the foundational conditions under which financial services can be expanded. Fiscal decentralization has a strong negative effect that can lead to an unbalanced distribution of finances because of the competition for resources. The values of openness to foreign trade and retail sales of goods are insignificant and have minor effects on finance.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 2: DID Baseline Regression Results

Variable	(1)	(2)	(3)
DID (opening dummy × early opening year)	12.87*** (3.29)	11.42*** (2.87)	9.36*** (2.41)
Population size		21.35*** (4.12)	18.72*** (3.65)
Population density			3.48** (1.72)
Economic density			2.79* (1.63)
Economic development level			4.21** (1.94)
Financial development level			5.83*** (2.11)
Openness to foreign trade			1.66 (1.28)
Urbanization rate			7.42*** (2.33)
Fiscal decentralization			-2.18* (1.29)
Industrial structure upgrading			3.97** (1.81)
Human capital level			2.55* (1.47)
Retail sales of commodities			1.83 (1.35)
Observations	12,863	12,863	12,863
Adjusted R^2	0.124	0.187	0.251
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Clustered standard errors	City	City	City
DID (opening dummy × early opening year)	12.87*** (3.29)	11.42*** (2.87)	9.36*** (2.41)

*Note: *** , ** and * stand for significance at 1, 5, and 10 percent levels, respectively; values of t-are given in brackets.*

5.1.2 PSM-DID Regression Results

Table 3 provides the outcome of the PSM-DID regression with controlled variables added in sequence. The results without control variables are presented in column (1), the DID coefficient is 8.74 and statistically significant at the 1 percent level. With the addition of population size in column (2), the DID coefficient is 7.96 and still very significant. With the inclusion of all control variables in column (3), the DID coefficient is 7.18, and the results continue to be robust.

Table 3: PSM-DID Baseline Regression Results

Variable	(1)	(2)	(3)
DID (opening dummy × early opening year)	8.74*** (3.98)	7.96*** (3.43)	7.18*** (3.02)
Population size		18.59*** (3.95)	16.28*** (3.48)
Population density			3.11** (1.67)
Economic density			2.47* (1.51)
Economic development level			3.81** (1.84)
Financial development level			5.35*** (2.06)
Openness to foreign trade			1.38 (1.17)
Urbanization rate			6.85*** (2.24)
Fiscal decentralization			-1.93* (1.15)
Industrial structure upgrading			3.48** (1.71)
Human capital level			2.26* (1.32)
Retail sales of commodities			1.54 (—)
Observations	9,125	9,125	9,125
Adjusted R^2	0.285	0.352	0.426
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Clustered standard errors	City	City	City

*Note: * , ** and *** indicate significance at the 1, 5 and 10 percent level, respectively, and t-values are provided in parentheses.*

Table 4 shows the average treatment effect on the treated (ATT) after PSM matching. The impact of high-speed rail opening on the overall number of bank branches is positive and statistically significant at the 1 percent level (t-value = 3.98), i.e., it is 8.74. It is in accordance with the results of the DID baseline regression, which means no interference caused by the bias of sample selection remains and the drawn conclusion is credible.

Table 4: Average Treatment Effect on the Treated (ATT) for PSM-DID

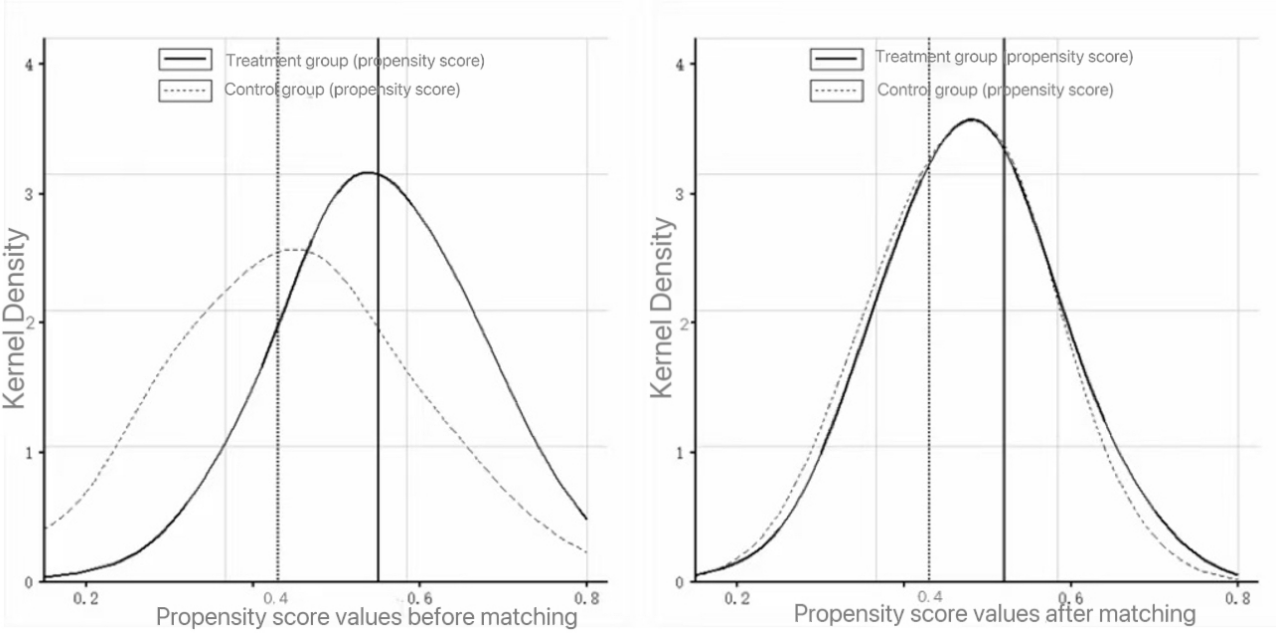
Variable	ATT	Std. Err.	t-value
Total number of bank branches	8.74***	2.19	3.98

*Note: ***, ** and * are used to indicate significance at the 1 percent, 5 percent and 10 percent level respectively.*

© 2026 The Author(s). Authors retain copyright in their articles.

The PSM-DID matching results are shown in Figure 1.

Figure 1: PSM-DID Matching Results

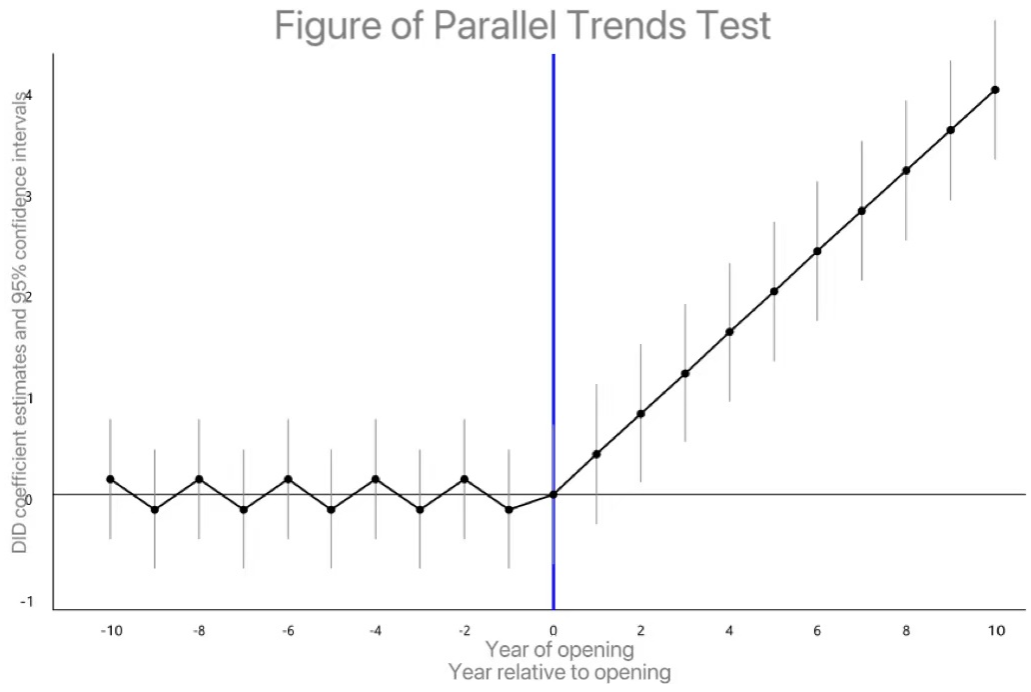


5.2 Robustness Checks

5.2.1 Parallel Trends Test

In order to check whether the difference-in-differences model is valid, a parallel trends test is carried out. The findings suggest that all pre-policy period coefficients are insignificant, thus it can be concluded that the treatment group and the control group meet the parallel trends assumption. Following the introduction of the policy, the coefficients are strongly positive and remain increasing with time, indicating that the policy impact of high-speed rail opening is real and emerging slowly, and the model specification is sensible and correct. Figure 2 shows the results of the parallel trends test.

Figure 2: Parallel Trends Test



© 2026 The Author(s). Authors retain copyright in their articles.

5.2.2 Replacing the Dependent Variable

In order to ensure that the main conclusions are not sensitive to the selection of indicators, natural logarithms of loan balances, deposit balances, and the number of bank branches are taken as dependent variables. The results of the regression indicate that the DID coefficients are 0.166, 0.154 and 0.198, respectively, and statistically significant at the 1 percent level of significance, as expected in the baseline results.

5.2.3 Excluding Special Samples

Local governments have unique economic and financial foundations. Having excluded them, the sample interference of special samples disappears and it can be tested whether the findings are generalizable. The DID coefficient is 8.63 after the exclusion of municipalities, which is significant at the 1 percent level, thus demonstrating that the base rate results are not compromised due to special sample interference and that the results are strong and believable. Table showing the result of robustness tests.

Table 5: Robustness Checks

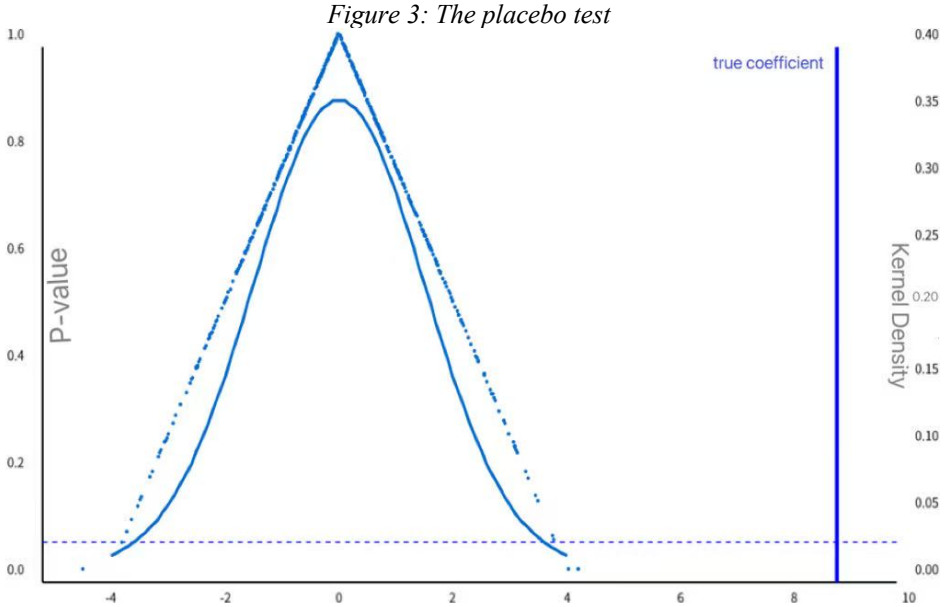
Variable	(1) Replacing dependent variable: ln loan balance	(2) Replacing dependent variable: ln deposit balance	(3) Replacing dependent variable: ln number of bank branches	(4) Excluding municipalities
DID	0.166*** (4.18)	0.154*** (4.11)	0.198*** (4.07)	8.63*** (3.85)
Control variables	Controlled	Controlled	Controlled	Controlled
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	8,562	8,562	8,562	7,914
Adjusted R^2	0.685	0.671	0.702	0.656

*Note: *** and **, and * represent the significance level of 1 percent, 5 percent and 10 percent respectively; t-values are given in brackets.*

5.2.4 Placebo Test

The findings indicate that the coefficients of the fictitious policy shocks are centered on zero whereas the real policy coefficient is quite far away from the random distribution which means that the effect of the high-speed rail opening policy can be real and not due to any random effects, thus, excluding interference with omitted variables and randomness. The results of the placebo test are illustrated in Figure 3.

Figure 3: The placebo test



5.3 Heterogeneity Analysis

Heterogeneity is analyzed from three dimensions: region, city size, and administrative level.

© 2026 The Author(s). Authors retain copyright in their articles.

5.3.1 Regional Heterogeneity

Table 6 displays the outcomes. By geographic area the coefficients of the DID for the eastern region are 10.82, which is statistically significant at the 1% level; the coefficients of the central and western regions are 3.69 and not significant. It means that the impact of the high-speed rail opening on financial resources distribution is stronger in the eastern region. This can be explained by the fact that the eastern region, due to the higher level of development of financial markets and the higher effectiveness of factor mobility, is more capable of taking advantage of the time-space compression dividends facilitated by high-speed rail, whereas the policy effect in the central and western regions has not been completely achieved yet.

5.3.2 Heterogeneity by City Size

As per the city size grouping, the DID coefficient of large cities is 9.57 which is significant at the 1 percent probability level, while the coefficient of small and medium-sized cities is 4.18 and is not significant. It implies that the impact of high-speed rail on financial resources is stronger in large cities. Large cities also have a greater density of financial institutions and corporate resources and can better attract and focus financial resources when the high-speed rail opens. Conversely, the basic conditions of small and medium-sized cities are relatively poor, and thus this driving effect is less visible.

5.3.3 Heterogeneity by Administrative Level

The grouped by administrative level DID coefficient of provincial capitals/central cities is 9.21, which is significantly positive at the 1 percent level; the coefficient of ordinary prefecture-level cities is 4.05 and insignificant. It indicates that financial resource allocation of provincial capitals and central cities is influenced by the opening of high-speed rail. Since they are the financial and economic centers of the region, these cities have better policy resources and financial service systems, and the high-speed rail strengthens this benefit. The financial attractiveness of ordinary prefecture-level cities is relatively weak, so the effect is not apparent.

Table 6: Heterogeneity Tests

Variable	(1) Regional heterogeneity - Eastern	(2) Regional heterogeneity - Central and Western	(3) City size heterogeneity - Large cities	(4) City size heterogeneity - Small and medium-sized cities	(5) Administrative level heterogeneity - Provincial capitals/Central cities	(6) Administrative level heterogeneity - Ordinary prefecture-level cities
DID	10.82***	3.69	9.57***	4.18	9.21***	4.05
	(3.41)	(1.57)	(3.08)	(1.69)	(3.15)	(1.73)
Controls	Controlled	Controlled	Controlled	Controlled	Controlled	Controlled
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1386	2902	1924	2364	1204	3082
Adjusted R^2	0.708	0.593	0.691	0.617	0.695	0.613

6. Research Conclusions and Policy Recommendations

6.1 Research Conclusions

This study uses DID and PSM-DID techniques in analyzing the effects of high-speed rail (HSR) opening on the financial development of regions based on panel data of Chinese cities between 2007-2022 by using the total number of bank outlets and the balances of deposits and loans of financial institutions as proxy indicators. The reliability of the results of the research is guaranteed by means of parallel trend tests, robustness checks, and placebo tests. The results show that the opening of HSR has a significant and enduring positive empowerment impact on regional financial development. Not only does it increase the geographical reach of financial services, but it also increases the extent of regional capital availability and agglomeration, which acts as a valuable exogenous push factor in encouraging regional financial development. Both policy shocks and foundational conditions contribute to the development of regional financial development. Economic growth,

© 2026 The Author(s). Authors retain copyright in their articles.

population agglomeration, financial deepening, advancements in urbanisation, industrial structure upgrading, and human capital enhancement are all vital drivers of financial development. Nevertheless, there is some degree of counterbalance to the balanced distribution of financial resources since fiscal decentralisation causes inter-regional resource competition.

The financial empowerment consequences of HSR opening are highly context-specific. With regards to city size, large and medium-sized cities, because of their larger population-economic agglomeration and stronger financial conditions, have much more capacity to absorb the dividends of factor flows, and the effect is therefore much higher than in small cities. Geographically speaking, the eastern region, which has a strong economic foundation and an active financial market, demonstrates a stronger agglomeration-strengthening effect of HSR on finance. Despite the fact that the central and western regions also enjoy enhanced factor mobility, the impact remains weaker because the industrial base and human capital remain limited. In line with the administrative hierarchy, the opening of HSR is more significant in terms of financial resource allocation in provincial capitals and central cities. They are the most complete centres of financial and economic activities in the region, and HSR contributes to them further. On the contrary, the financial attractiveness of ordinary prefecture-level cities is relatively weaker, and it is not as obvious. Also, parallel trends and placebo experiments support the idea that there is truly a causal connection between the opening of HSR and the development of finances at the level of the region, and the empowerment effect cannot be considered as resulting due to irrelevant factors like time trends or any other policy shock.

The HSR opening eliminates geographical boundaries, facilitates the movement of production factors between regions, and adds a new impulse to regional financial development. The scale at which this financial empowerment effect will be achieved depends directly on the regional developmental conditions of the foundation, hence offering an empirical proof of how coordinated financial development can be attained by optimizing the HSR network structure and enhancing the regional developmental bases.

6.2 Policy Recommendations

Using the results of the research, the following recommendations are suggested in order to use the positive aspects of HSR even more intensively in the promotion of the regional financial development, to gradually enhance and equalize the financial status of the regions, to reduce differences in the inter-regional financial development, and to create a more just and effective distribution of the financial means.

Initially, the layout of the HSR network must be optimized. It is necessary to ensure the accessibility and convenience of transport-deficient zones based on the travel needs of the residents in small and medium-sized cities, and the central and western regions. The transportation situation ought to be consistently made better in order to minimize time disparities as well as spatial space between cities and allow further underdeveloped areas to use the benefits of HSR to attract financial assets, thus establishing the basis of balanced regional financial development. The next step would be to reinforce the conditions of urban development, particularly in small and medium-sized cities. Continuous improvement of industrial environments, public services, and other forms of infrastructure should make cities more attractive in general. With a stronger development base alone, funds will be able to be kept in the city and financial institutions will be lured into opening branches so that the development potential offered by HSR does not just stop at the stage of better transport conditions.

Moreover, the concept of differential development strategies should be followed depending upon the peculiarities of the place and should not use a universal solution. Eastern region and major cities ought to utilize the benefits of HSR to develop the financial service capacity and innovation potential, and big and middle cities in the central and western regions need to reduce their weaknesses step-by-step depending on their specifics, constantly increase the financial development level, and gradually reduce the gaps between them and more developed areas in financial development. At the same time, unreasonable inter-regional competition should be minimized, and inter-regional coordination and cooperation enhanced to remove limitations created by administrative barriers. Finance needs to be steered to become more diversified, which would enhance the effectiveness of total capital use and make it easier to serve local economic growth and thus realize the supplementary and mutual reinforcement objectives of transport development and financial development. Also, once the HSR is opened, monitoring of various trends, including capital flows, institutional layouts, and credit scales, should continue to be dynamically monitored. The financial operation safety bottom

© 2026 The Author(s). Authors retain copyright in their articles.

line in financial development promotion should be strictly adhered to, the risk prevention mechanism enhanced, and the stability, health, and sustainable development of regional finance guaranteed.

References

[1] Yansui Liu. (2018). Research on the Urban-rural Integration and Rural Revitalization in the New Era in China. Acta Geographica Sinica, 73 (4), 637–650.

[2] Yufei Wang, Pengfei Ni. (2016). Economic Growth Spillover and Spatial Optimization of High-speed Railway. China Industrial Economics, (2), 21–36.

[3] Xiujuan LAN, Weiguo Zhang. (2023). Economic Linkage Pattern of Urban Clusters and Coordinated Regional Economic Development under the Influence of High-speed Rail Network. Statistics & Information Forum, (10), 21-34.

[4] Yalei Yan, Tao Yu, Jianghao Lai. (2022). the Impact of High Speed Rail on Regional Accessibility and Spatial Linkage Patterns: A Case Study of the Yangtze River Delta City Cluster. Shanghai Urban Planning Review, (6), 153–158.

[5] Hongchang Li. (2019). the Economic Gradient Effect of High-speed Rail on Small and Medium-sized Cities and Development Countermeasures. Railway Economics Research, (3), 1–5.

[6] Fangzhi Fan, & Lijun Zhang. (2003). a Study on the Transformation of Regional Financial Structure and Upgrading of Industrial Structure in China. Journal of Financial Research, (11), 36–48.

[7] Li Zhou. (2002). an Empirical Analysis of Regional Financial Development and Economic Growth in China: 1978–2000. Journal of Financial Research, (10), 1–13.

[8] Jian Zhao. (2025). Digital Finance and High-quality Development of Manufacturing: The Driving Path of Innovation Factor Agglomeration and Resource Allocation Efficiency. Science and Technology Management Research, (13), 155-166.

[9] Jiawen Cheng, Hui Hui Cui. (2025). High-speed Rail Opening, Financial Agglomeration, and House Price Expectations. Construction Economy, (1), 34-43.

[10] Ying Wu, Xianping Zhou, & Benyan Tan. (2021). High-speed Rail Opening, Bank Competition and Corporate Debt Financing Costs. Finance Forum, 26 (11), 27–36.

[11] Xin Li, Fuping Li. (2021). Research on the Impact of High-speed Rail Opening on the Development of Digital Inclusive Finance: Evidence from a Quasi-natural Experiment of Prefecture-level Cities. Industrial Technology & Economy, 40 (11).

[12] Perroux, F. (1950). Economic Space: Theory and Applications. the Quarterly Journal of Economics, 64 (1), 89–104.

[13] Huaxing Wang, Daqian Shi, & Hongwei Yu. (2019). CAN High-speed Rail Opening Promote Coordinated Regional Economic Development?. Shanghai Journal of Economics, (11), 59–69.

[14] Fujita, M., Krugman, P., & Venables, a. J. (1999). the Spatial Economy: Cities, Regions, and International Trade. MIT Press.

[15] Qin, Y. (2017). “NO County Left Behind?” the Distributional Impact of High-speed Rail Upgrades in China. Journal of Economic Geography, 17 (3), 489–520.

[16] Friedmann, J. (1966). Regional Development Policy: A Case Study of Venezuela. MIT Press.

[17] Clark, G. L., Feldman, M. P., Gertler, M. S., & WóJcik, D. (Eds.). (2007). the New Oxford Handbook of Economic Geography. Oxford University Press.

[18] Central Committee of the Communist Party of China and the State Council. (2025). Opinions of the Central Committee of the Communist Party of China and the State Council on further deepening rural reform and solidly promoting the comprehensive revitalization of rural areas. Gov.cn. https://www.gov.cn/gongbao/2025/issue_11906/202503/content_7011166.htm.

© 2026 The Author(s). Authors retain copyright in their articles.

[19] National Railway Administration. (2025). Railway Statistical Bulletin 2024. National Railway Administration. https://www.nra.gov.cn.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.4752
Published by: Zeus Press

Analysis of the Impact of Financial Technology on the Liquidity Risk of Commercial Banks

Mingrui Su*

School of Finance, Jilin University of Finance and Economics, Changchun, China

**Corresponding author: Mingrui Su.*

Abstract

With the rapid development of financial technology, the liquidity of commercial banks has undergone tremendous changes. Based on this, this paper comprehensively combs through a number of relevant core literature at home and abroad, discusses the impact of financial technology on the liquidity risk of commercial banks, the mechanism of financial technology on the liquidity of commercial banks, summarizes relevant research conclusions and puts forward corresponding suggestions. The results show that most scholars believe that fintech can increase the liquidity of commercial banks, improve the level of liquidity creation and reduce risks. The mechanism includes the effect of debt structure optimization and the effect of asset allocation efficiency improvement. These mechanisms reshape the liquidity and risk flow of commercial banks. A few scholars believe that the development of financial technology will bring liquidity challenges to commercial banks. The research conclusion provides a decision-making reference for the management of liquidity risk of commercial banks and the regulatory authorities to improve the supervision of financial technology.

Keywords

fintech, commercial banks, liquidity risk

1. Introduction

Finance is the blood of the national economy and an important part of the country's core competitiveness. The report of the 19th National Congress of the Communist Party of China proposed to “enhance the ability of financial services to the real economy”. In 2018, China issued the measures for liquidity risk management of commercial banks, emphasizing the need to strengthen liquidity risk management of commercial banks and ensure the safe and stable operation of the banking system. In the context of the digital economy, the rapid development of financial technology has profoundly reshaped the financial ecosystem and pattern through big data, artificial intelligence, cloud computing, blockchain and other technologies. The deep integration of fintech and commercial banks has constantly created new areas of finance. In the context of the impact of financial technology, traditional commercial banks' business operations are difficult to adapt to the digital management architecture, and the liquidity problem is prominent [1]. At the academic level, on the one hand, the existing literature mostly demonstrates the correlation between the two separately and lacks a systematic literature review. For example, the development of financial technology has significantly increased the liquidity risk of commercial banks on the whole, and the impact is more obvious in urban commercial banks [2]. On the other hand, the existing literature on business segmentation is insufficient [3]. Therefore, the study

© 2026 The Author(s). Authors retain copyright in their articles.

of financial technology on the liquidity risk of commercial banks and how to improve the liquidity ability of commercial banks has important theoretical and practical significance for deepening the reform of the financial system.

This paper systematically combs through the relevant core literature, clarifies the transmission path of the impact of financial technology on the liquidity risk of commercial banks, explores the mechanism of the impact of financial technology on bank liquidity, and expands the trend of liquidity risk management and risk prevention boundary of commercial banks. The research conclusion enriches the theoretical research on bank liquidity risk, and provides an empirical reference for commercial banks to better use fintech to manage liquidity risk and serve the real economy.

2. Current Situation and Background of FinTech and Liquidity Management of Commercial Banks

Under the background of the rapid development of financial technology, the liquidity of domestic commercial banks shows obvious term mismatch and short-term deposit characteristics, which increases the liquidity pressure of large, medium and small banks. With the vigorous development of financial technology, commercial banks continue to innovate the financial service mode through the deep integration of emerging technologies such as big data and intelligent liability management. Some scholars believe that it reduces the liquidity creation function of banks, but mainly reduces the liquidity of the liability side, and has no significant impact on the liquidity of the asset side [4]. In the process of liquidity creation of commercial banks, it is closely related to the development of the real economy, and the two affect each other [5]. On the whole, the structural characteristics of the liquidity risk of commercial banks at this stage are obvious: the liquidity pressure of large banks is small, and small and medium-sized banks are more vulnerable to the liquidity impact caused by capital fluctuations [4]. From the perspective of assets and liabilities, fintech can improve the liquidity creation of commercial banks in general [6]. Under the influence of financial technology, Internet deposits have changed the debt structure of banks and profoundly changed the original liquidity pattern of commercial banks [7].

In view of the limitations of the original liquidity management theory, combing the relevant literature, it is found that the academic community has supplemented and revised the theoretical framework. First, the debt side theory is updated. Internet Finance and third-party payment continue to divert low-cost bank deposits, resulting in a significant decline in the stability of bank deposits, and the passive debt model is no longer applicable. Big data technology can accurately predict customers' capital behavior and form the digital liability management theory [8]. Banks use algorithms to predict deposit fluctuations and allocate funds in advance. Second, the iteration of asset side theory. Relying on big data risk control, commercial banks' online microfinance can achieve small, decentralized, high-frequency, improve asset liquidity, ease the liquidity pressure brought by long-term credit, and enrich the application boundary of expected income theory [9]. Third, the expansion of risk contagion theory. The traditional theory only considers the liquidity of a single bank. In the digital era, the cross-platform flow speed of funds has increased significantly, and the local liquidity tension will spread rapidly. Scholars have incorporated the network infection mechanism into the liquidity theory, forming a systematic liquidity analysis framework in the digital environment. In general, the theory of liquidity management has moved from single asset management to asset liability coordination. Under the influence of financial technology, it has added new content such as digital prediction and network risk contagion. The theoretical system has been constantly adapted to the current real environment of the banking industry.

3. Factors Affecting the Liquidity of Commercial Banks

3.1 Influencing Factors

The liquidity of commercial banks will be affected by both traditional factors and financial technology. Traditional influences include macroeconomic policies, economic cycles, banking regulation, and residents' savings habits. The results show that economic policy uncertainty (EPU) is the key transmission channel for monetary policy to affect banking output, and it is still strong in the indicators of bank liquidity creation [10]. The implementation of relevant regulations has significantly enhanced the liquidity creation of traditional deposit and loan business [11]. Other studies believe that reasonable development and promotion of green

© 2026 The Author(s). Authors retain copyright in their articles.

credit business will improve the liquidity of bank funds. However, due to the promotion of economic globalization, commercial banks frequently carry out interbank business and increase their dependence on the interbank market; their liquidity risk has become increasingly prominent. Therefore, among the factors affecting the liquidity of commercial banks, traditional factors still play an important role. As an emerging technical force, fintech injects new vitality and challenges into the business and liquidity of banks.

In terms of the impact of fintech, most scholars believe that the technology spillover effect of fintech is greater than the market crowding out effect, which can improve the liquidity creation ability of commercial banks. On the asset side, commercial banks use technology updates to carry out online terminal services, accurately capture users' capital behavior, timely promote business products, and achieve the ability to quickly attract customers. Fintech promotes liquidity creation through the technology spillover effect. On the liability side, fintech inhibits liquidity creation through the market crowding out effect, thereby improving the efficiency and risk management level of financial service institutions [6]. In addition, the development of digital finance in favor of financial formats can further slow down the liquidity hoarding of commercial banks and make users' financial services more personalized, transparent and convenient. Fintech promotes technological innovation and improves liquidity by reducing enterprise financing costs. Fintech has a positive drive on business innovation and technological progress of commercial banks.

4. The Mechanism of the Impact of FinTech on the Liquidity of Commercial Banks

4.1 Effect of Debt Structure Optimization

The liquidity risk of banks stems from the maturity mismatch of assets and liabilities and risk transmission, which is specifically manifested in the lack of liquidity at the asset side or the sudden rise of financing costs at the liability side, and its risk-taking level is affected by risk appetite and external shocks. However, the in-depth application of financial technology reshaped the structure of bank liabilities: the allocation of residents' financial assets continued to tilt towards financial products, the deposit business was under pressure, and it was more difficult for banks to obtain low-cost funds, forcing banks to optimize the matching of assets and liabilities and improve the liquidity early warning mechanism. This phenomenon is affected by the dual role of the development of financial technology inside and outside the bank. At the level of debt structure optimization, fintech relies on big data technology to accurately obtain customers, break geographical restrictions through online channels, broaden deposit sources, improve their operational efficiency through business process optimization, strengthen risk management with the help of intelligent risk control, and improve the cost structure by using scientific and technological means [12]. Efficient online settlement and other systems accelerate capital turnover, increase the proportion of core stable deposits, reduce the volatility of deposits, consolidate the stability of the bank's liability side, and improve the overall liquidity level [13].

The development of banking financial technology will increase the ratio of non-deposit liabilities to deposit liabilities, reduce the relative proportion of liabilities-side deposits, and increase the proportion of demand deposits in deposits, which will affect the liquidity risk tolerance of banks. Through the above means, we can reshape the liquidity operation logic of commercial banks and increase the proportion of highly liquid assets. However, the development of non-banking financial technology industry has led to the loss of bank deposits and increased the difficulty of daytime liquidity management. Most of the existing literature focuses on the analysis of the external impact of financial technology, and there are differences in the research conclusions, such as the views that Internet finance promotes bank risk, the scale of third-party payment and the liquidity of large banks have heterogeneous correlation [14].

4.2 Effect of Improving Asset Allocation Efficiency

Relying on financial technology, it has brought more financing channels for commercial banks in terms of resource allocation. Through the Internet, blockchain and other technologies, banks can more easily obtain funds and flexibly adjust their asset portfolios, reduce financing costs and improve financing efficiency. In the business process, the digital settlement system simplifies the transfer, receipt and payment process, and accelerates the social capital turnover. The big data and intelligent risk control model can accurately identify the relevant risk coefficient of the borrower, shorten the credit approval cycle, and speed up the flow of credit assets. And because of the application of advanced technology, commercial banks have realized intelligent

© 2026 The Author(s). Authors retain copyright in their articles.

business processes, which not only improves the speed and accuracy of bank business processing, but also reduces the cost in the process of resource allocation.

The development of fintech has expanded the scope of factor market allocation, eliminated hidden and institutional barriers that hinder the free flow of factors and orderly market access, and improved the overall factor productivity of real enterprises. It has further improved the price formation and adjustment mechanism and effectively corrected the problems existing in traditional finance, such as “resource mismatch”, “credit mismatch” and attribute, industry and stage-related mismatches. By improving the allocation of resources, promote the optimization of internal governance of commercial banks and further improves the efficiency of factor allocation. The further concentration of capital has attracted the convergence of high-end talents and technologies, fully responding to the core requirements of the development of new quality productivity and meeting its rigid demand for high-skilled, digital and advanced production factors. This ultimately promoted the improvement of market resource allocation efficiency and overall economic efficiency [15].

4.3 Long Tail Effect

The long tail effect was proposed by chrisanderson, editor-in-chief of Wired magazine, in 2004. It means that by focusing on a large number of niche markets or low-volume products, the overall market value will be greater than that of the mainstream market. Its core is to meet the huge niche demand. Although the profit of a single product is low, the overall market potential is huge. For commercial banks, long tail funds are small in size and large in quantity, sensitive to some customers, and have special needs, which are easy to disturb the liquidity of commercial banks. The development of new businesses and products with the help of fintech can promote personalized service and product diversification, and then attract more financial resources to match many long-tail customers and expand the bank's liquidity coverage. At the same time, the push and service mechanism of fintech superimposed on the digital system will break the edge effect, cover more small and Micro customers and meet the diversified new needs of customers, promote more funds to participate in the creation of liquidity in commercial banks, and improve the overall liquidity creation level of commercial banks.

5. Trends and Risks of Bank Liquidity Management under the Impact of FinTech

5.1 Development Trend

Under the background of the full penetration of financial technology, digital risk control has become a normal means of bank operation. The traditional liquidity management mode relies on the big data platform to realize the dynamic real-time monitoring of funds, and greatly improve the management accuracy. There is an effective threshold between the frequency of fintech risk control review and the intensity of punishment. Appropriate external supervision can form a positive cycle of “supervision → improvement → transition”. The forced back effect of risk thunderstorm events can accelerate the adoption and governance upgrading of fintech [16]. In addition, the risk level of city banks' credit is reduced. City commercial banks can manage deposit and loan business more efficiently, reduce the deposit-to-loan ratio, have more liquidity reserves, and better respond to potential liquidity shocks, thereby reducing the overall risk level [17].

5.2 Facing Challenges

With the increasing dependence of financial business on artificial intelligence, blockchain and other technologies, problems such as algorithm black box, data leakage and system vulnerabilities have become increasingly prominent. Some scholars believe that due to the convergence of product and service modes (homogeneous competition), differentiated competitive advantages have not been effectively formed [18]. Massive customer data brings a hidden danger of data leakage, and the algorithm model may have logical defects, which is easy to form model risk. Some scholars also believe that financial technology innovation improves bank risk, such as increasing risk exposure through product innovation and business expansion, and the impact of digital Finance on the financial environment, such as the process in which transaction costs and cognitive biases are integrated into systematic results, leading to risk spillovers [19]. In addition, the increase in the intensity of financial regulation will inhibit the impact of digital Finance on the risk-taking level of commercial banks, and it is very easy to induce systemic liquidity risk [20].

© 2026 The Author(s). Authors retain copyright in their articles.

6. Conclusion

Based on the existing literature, it is concluded that in recent years, scholars have mostly studied the correlation between financial technology and liquidity of commercial banks, and less systematic combing of the two. Second, the research on the liquidity creation mechanism of commercial banks by fintech and the relevant literature are less. Third, since entering the "Fifteenth Five-Year Plan", China has carried out less research and application of fintech in the liquidity of commercial banks. Fourth, on the whole, fintech can significantly improve the liquidity creation level of commercial banks.

From the bank level, commercial banks should take the initiative to lay out the construction of financial technology, use big data and other technologies to accurately predict the capital flow, stabilize the debt scale, alleviate the liquidity pressure caused by short-term deposits, optimize the credit asset structure with the help of digital risk control, and reduce the maturity mismatch. It is committed to building a new credit risk control system driven by science and technology, improving the ability of intelligent risk control, ensuring the quality of business and the safety of credit funds, and promoting the in-depth transformation of commercial banks from traditional credit mode to digitalization and intelligence. Moreover, commercial banks should implement differentiation strategies according to their own size, market demand and business scale to continuously improve their competitiveness. For the government regulatory departments, they should maintain the sound operation of the banking industry, establish a dynamic liquidity monitoring system, supplement the weakness of the supervision of Internet financial business, guide banks to develop digital business rationally, prevent disorderly competition, and hold the bottom line of the liquidity safety of the banking system, to prevent the occurrence of systemic risks.

References

[1] Morelli, G., Pugliese, V., & Santucci de Magistris, P. (2025). Liquidity coverage at risk. Quantitative Finance, 25(2), 291–306. https://doi.org/10.1080/14697688.2025.2457366

[2] Cheng, L. (2026). The impact of financial technology on the liquidity risk of commercial banks. Business Observation, 12(12), 69–72.

[3] Yin, F., Jiao, X., Zhou, J., Yin, X., Ibeke, E., Iwendi, M. G., & Biamba, C. (2022). Fintech application on banking stability using big data of an emerging economy. Journal of Cloud Computing, 11(1). https://doi.org/10.1186/S13677-022-00320-7

[4] Li, J., Wang, L., & Peng, Y. (2023). Bank fintech and liquidity creation. Nankai Economic Studies, (5), 3–18. https://doi.org/10.14116/j.nkes.2023.05.001

[5] Berger, A. N., & Sedunov, J. (2017). Bank liquidity creation and real economic output. Journal of Banking & Finance, 81, 1–19.

[6] Zhao, X., Li, S., Lu, K., & Zhong, Y. (2024). Intelligent transformation, fintech, and green growth: A general equilibrium analysis based on credit allocation perspective. Journal of Environmental.

[7] Guo, P., & Zhang, C. (2023). The impact of bank fintech on liquidity creation: Evidence from China. Research in International Business and Finance, 64. https://doi.org/10.1016/J.RIBAF.2022.101858

[8] Yu, W., & Zheng, J. (2021). Internet finance, deposit competition, and bank liquidity creation. West Forum, 31(2), 106–124.

[9] Julapa, J., Lauren, L., & Timothy, L. (2021). Fintech lending and mortgage credit access. The Journal of FinTech.

[10] Lee, E., & Alam, M. Z. (2026). Impact of monetary policy on bank liquidity creation: Economic policy uncertainty channel. Journal of Banking Regulation, 27(3), 39–39. https://doi.org/10.1057/S41261-026-00340-2

[11] Li, M., Zhang, H., & Ruan, P. (2026). Strong financial supervision and divergence in bank liquidity creation: Empirical evidence from new asset management regulations. Finance Research Letters, 107, 109888.

© 2026 The Author(s). Authors retain copyright in their articles.

[12] Wu, X. (2022). Financial technology under the background of China Merchants Bank retail business transformation analysis. Industrial Engineering and Innovation Management, 5(12). https://doi.org/10.23977/IEIM.2022.051204

[13] Song, K., Li, Z., & Yang, J. (2023). Fintech and bank behavior: From the perspective of liquidity creation. Journal of Financial Research, (2), 60–77.

[14] Liu, Y. (2026). Research on the liquidity of smart deposits through financial technology innovation under the framework of asset liability management. Procedia Computer Science, 282, 2023–2031. https://doi.org/10.1016/J.PROCS.2026.05.245

[15] Tang, Q., Xu, K., Zou, M., & Huo, L. (2026). Can FinTech empower companies to develop corporate new quality productive forces? Evidence from Chinese listed companies. International Review of Economics & Finance, 108, 105298. https://doi.org/10.1016/J.IREF.2026.105298

[16] Li, G., Huang, Y., & Wang, Q. (n.d.). Four-party evolutionary game analysis of collaborative governance of digital supply chain financial risks embedded with fintech. Systems Engineering, 1–13.

[17] Liu, Y., Tong, A., & Nourani, P. (2026). FinTech adoption and credit risk dynamics: Evidence from China's A-share banking sector. Journal of Banking Regulation, 27(2), 25–25. https://doi.org/10.1057/S41261-026-00327-Z

[18] Ren, M., Zhao, J., Ke, K., & Li, Y. (2023). Bank homogeneity and risk-taking: Evidence from China. The Quarterly Review of Economics and Finance, 92, 142–154. https://doi.org/10.1016/J.QREF.2023.09.002

[19] Nofsinger, J. R. (2026). FinTech as behavioral infrastructure: Digital platforms and the amplification of risk. Journal of Behavioral and Experimental Finance, 51, 101216. https://doi.org/10.1016/J.JBEF.2026.101216

[20] Wang, R., & Cui, W. (2024). Research on digital finance, financial regulation and commercial bank risk-taking. Economics & Management Review, 5(3). https://doi.org/10.37420/J.EMR.2024.046

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.5360
Published by: Zeus Press

Tax Incentives and Debt-Financing Efficiency in Advanced Manufacturing Enterprises: Evidence from Chinese A-Share Listed Firms

Pengfei Qian*

University of Shanghai for Science and Technology, Shanghai, China

**Corresponding author: Pengfei Qian.*

Abstract

Advanced manufacturing firms face long investment cycles, high technological uncertainty, and limited pledgeable assets, so the efficiency with which they use debt is central to sustained innovation. This study examines the association between tax incentives and debt-financing efficiency among Chinese A-share listed advanced manufacturing firms. Annual efficiency scores are estimated using an input-oriented CCR data envelopment analysis model, and two-limit Tobit regressions are applied to 16,817 firm-year observations from 2015 to 2025. Tax incentives are measured as tax refunds received relative to operating revenue. The fully controlled coefficient is 0.5549 and the average marginal effect is 0.5039, indicating a significant positive association. Financing-constraint relief accounts for approximately 15.09% of the total association. The relationship is significantly stronger for non-state-owned and high-technology firms, whereas regional and firm-size differences are not statistically significant. Lagged and re-winsorized estimates support the baseline result, but alternative definitions of incentives and efficiency reverse the sign, revealing material measurement sensitivity. The findings suggest that timely tax relief can improve debt allocation, although interpretation should remain associational and conditional on the selected measures.

Keywords

tax incentives, advanced manufacturing, debt-financing efficiency, data envelopment analysis, financing constraints

1. Introduction

Advanced manufacturing is distinguished from conventional scale-oriented production by its dependence on specialized equipment, technical knowledge, continuous process improvement, and repeated investment before commercial returns become stable. Debt remains a major source of external capital because bank credit and other debt instruments can be obtained without diluting ownership, but the terms on which debt is supplied may not match the economic life of innovative projects.

Under information asymmetry, lenders may ration credit rather than simply raise the interest rate [1], while firms with valuable projects may avoid external finance when managers believe that outside investors cannot observe project quality [2]. A firm can consequently carry substantial debt while still financing inefficiently if

© 2026 The Author(s). Authors retain copyright in their articles.

it pays too much, assumes excessive balance-sheet risk, or fails to convert borrowed funds into adequate operating returns.

A refund received during the investment stage retains cash inside the firm and reduces the need for emergency borrowing. A positive effect on financing availability therefore does not necessarily imply a positive effect on financing efficiency.

Capital-structure theory shows that corporate taxation changes the relative cost of debt and equity [3], and recent research links tax relief to financing constraints and investment allocation [4][5]. Yet relatively few studies ask whether tax incentives improve the joint relationship among debt scale, financing cost, leverage capacity, and capital return in advanced manufacturing.

This study uses a two-stage framework and then tests timing, measurement sensitivity, financing-constraint mediation, and heterogeneity across region, size, ownership, and technology groups. It shifts the policy question from whether firms receive more funds to whether debt inputs are matched by financing capacity and capital return.

2. Theoretical Analysis and Hypotheses

2.1 Direct Efficiency Channel

In the classic corporate-tax extension of capital-structure theory, the deductibility of interest affects the relative attraction of debt. A tax refund creates an immediate or near-term cash benefit that can substitute for part of external finance and can alter the timing, size, and maturity of new borrowing. For a firm facing a temporary liquidity gap, this benefit may prevent the use of expensive short-term credit.

Better liquidity may also allow the firm to negotiate longer maturities that correspond to the payoff period of equipment or research projects. In the input-oriented DEA setting used below, such a firm moves closer to the best-practice frontier because it can maintain outputs while proportionally compressing financial inputs.

The signal is supplementary rather than conclusive: public support does not guarantee commercial success, and creditors must still assess repayment capacity. This reasoning yields the first hypothesis: H1. Tax incentives are positively associated with the debt-financing efficiency of advanced manufacturing enterprises.

2.2 Financing-constraint Channel

Financing constraints arise when firms cannot fund all projects with positive expected value at an appropriate price and maturity. When the feasible financing set is narrow, managers may accept high-cost debt, roll over short-term loans repeatedly, postpone productive investment, or combine projects with financing instruments that have incompatible maturities.

Tax relief directly supplies internal liquidity and reduces the funding gap. The firm can then compare more financing alternatives, avoid distress-driven borrowing, and preserve investment with higher expected returns. Governance quality, project selection, and demand conditions still determine the final allocation, so financing-constraint relief should explain only part of the overall association.

Following Hadlock and Pierce [6], a larger ASA value in the coding used here denotes weaker financing constraints. If the direct Tax coefficient remains positive but declines after ASA enters the outcome equation, the results are consistent with partial mediation. H2. Tax incentives are positively associated with debt-financing efficiency partly through the alleviation of financing constraints.

3. Research Design

3.1 Data and Sample

The data combine CSMAR records, annual reports, and stock-exchange disclosures for firms classified as advanced manufacturing enterprises. Observations from 2014 are used only to construct lag-dependent values, including the average-asset and growth measures required for the 2015 observation; the estimation window is

© 2026 The Author(s). Authors retain copyright in their articles.

therefore 2015-2025. The estimation sample covers Shanghai and Shenzhen A-share firms and excludes Beijing Stock Exchange firms.

The cleaning procedure removes ST and *ST observations, pre-listing years, and firm-years with missing values for the main variables. It also excludes observations with non-positive total assets, operating revenue, or liabilities, as well as insolvent firms whose liabilities exceed assets. Continuous variables are winsorized at the 1st and 99th percentiles. After screening, the unbalanced panel contains 16,817 firm-year observations for 2,699 listed firms.

3.2 Variables and DEA Measurement

The dependent variable, debt-financing efficiency (Eff), is estimated separately for each year with the input-oriented CCR model developed by Charnes, Cooper, and Rhodes [7]. The model assumes constant returns to scale and asks how much the selected inputs could be proportionally reduced while the observed outputs are maintained. A score of one identifies a firm on the estimated annual frontier, whereas a score below one indicates proportional input-reduction potential relative to best-practice observations.

The DEA model uses two inputs and two outputs. The asset-liability ratio represents debt committed to the balance sheet, while debt-financing cost represents its price. Financial leverage and return on invested capital are outputs because they capture financing capacity and the return generated by the debt arrangement. Following related listed-company research [8], Table 1 reports the formulas and economic roles of these indicators. The model evaluates the observed combination relative to peers; it does not assume that greater leverage is unconditionally desirable. Potentially negative indicators are translated within year to satisfy non-negativity without changing their ranking.

Table 1: DEA input-output variable definitions

Category	Indicator	Measurement	Economic role
Input	Asset-liability ratio	Total liabilities / total assets	Debt scale
Input	Debt-financing cost	Interest expense / average short- and long-term interest-bearing debt	Debt price
Output	Financial leverage	EBIT / (EBIT - interest expense)	Debt-financing capacity and risk
Output	Return on invested capital	Net operating profit after tax / invested capital	Capital return

Note: Debt-financing cost uses the average balance of short- and long-term interest-bearing debt. Indicators that can be negative are translated within year before DEA estimation.

Tax equals tax refunds received divided by total operating revenue, which scales realized policy support to the firm's operating base. The baseline controls for Size, ROA, Profit, Growth, Fix, and Indep; Table 2 defines these variables. ASA, the absolute value of the SA index, enters the mediation test, with larger values denoting weaker financing constraints. Year indicators absorb shocks common to all firms, including broad changes in credit conditions and policy. The sensitivity analysis also scales refunds by total assets, but that denominator embeds capital intensity and is not economically equivalent to revenue scaling.

Table 2: Definitions of empirical variables

Category	Variable	Symbol	Definition
Dependent	Debt-financing efficiency	Eff	Annual input-oriented CCR DEA score based on Table 1; range 0-1
Explanatory	Tax incentives	Tax	Tax refunds received / total operating revenue
Mediator	Financing-constraint relief	ASA	Absolute value of the SA index; larger values indicate weaker constraints
Control	Firm size	Size	Natural logarithm of total assets
Control	Return on assets	ROA	Net profit / average total assets
Control	Profitability	Profit	Net profit / total operating revenue
Control	Growth	Growth	(Current-year operating revenue / prior-year operating revenue) - 1
Control	Fixed-asset ratio	Fix	Fixed assets / total assets
Control	Board independence	Indep	Independent directors / total directors x 100%

Note: ASA is used only in the mediation analysis. Indep is reported as a percentage.

© 2026 The Author(s). Authors retain copyright in their articles.

For firm o, the program chooses peer weights lambda and an input contraction factor theta. X and Y denote the annual input and output matrices, while x_o and y_o denote the evaluated observation. Theta is minimized subject to the weighted peers producing at least the observed outputs with no more than theta times the inputs. The maintained score lies between zero and one.

$$\min_{\theta,\lambda}\theta \text{s.t.} X\lambda \leq \theta x_0,\ Y\lambda \geq y_0,\ \lambda \geq 0 \tag{1}$$

3.3 Tobit Specification

The second stage relates Eff to Tax and the control variables. Ordinary least squares can predict values outside the feasible interval and does not explicitly accommodate the mass at the frontier. The baseline therefore uses a two-limit Tobit model [9]. The latent equation includes Tax, the control vector, year effects, and an idiosyncratic error. The observed variable equals zero when the latent value is below the lower limit, equals the latent value within the interval, and equals one when the latent value reaches the upper limit. Standard errors are clustered by firm to permit within-firm dependence over time.

Tobit coefficients describe the latent outcome rather than direct changes in observed efficiency, so the fully specified model is accompanied by an average marginal effect. Three models are estimated sequentially: Tax alone, Tax with firm controls, and Tax with controls and year effects. Stability across columns shows whether the result depends on observable composition or common annual shocks. The analysis remains associational because tax support can respond to firm conditions, managerial quality is unobserved, and DEA scores share a frontier estimated from the same sample.

A bounded outcome is not the only statistical issue in a two-stage DEA design. Simar and Wilson show that scores estimated from a common frontier are statistically dependent and recommend bootstrap truncated regression under their maintained assumptions [10]. Tobit is retained for comparability and because the available estimates use this design. Results are therefore interpreted as conditional associations; bootstrap frontier inference and exogenous policy variation remain extensions for future work.

$$\text{Eff}_{it}^{*} = \alpha_0 + \alpha_1 \text{Tax}_{it} + \beta' \text{Controls}_{it} + \gamma_t + \varepsilon_{it} \tag{2}$$

$$\text{Eff}_{it} = \max(0, \min(1, \text{Eff}_{it}^{*})) \tag{3}$$

4. Empirical Results

4.1 Descriptive Evidence and Diagnostic Checks

Table 3 reports the distribution of the estimation variables. Mean Eff is 0.2393 with a standard deviation of 0.1792, and only 40 observations reach the frontier value of one. The low mean does not imply that firms destroy value in an accounting sense; it indicates that, relative to the best performers in the same year and under the chosen inputs and outputs, most firms have substantial room to reduce debt exposure or financing cost while maintaining leverage and capital return. Tax has a mean of 0.0199, so refunds average about 1.99% of operating revenue. Its standard deviation of 0.0220 and range after winsorization show meaningful cross-firm and time variation.

Table 3: Descriptive statistics

Variable	N	Mean	SD	Min	Max
Eff	16,817	0.2393	0.1792	0.0430	1.0000
Tax	16,817	0.0199	0.0220	0.000005	0.1064
Size	16,817	22.2669	1.1685	20.1168	25.9272
ROA	16,817	0.0358	0.0665	-0.2177	0.2244
Profit	16,817	0.0480	0.1794	-0.9415	0.4107
Growth	16,817	0.1303	0.3105	-0.4953	1.6486
Fix	16,817	0.2101	0.1243	0.0142	0.5749
Indep	16,817	37.9842	5.3879	33.3300	57.1400

Note: Continuous variables are winsorized at the 1st and 99th percentiles. Indep is expressed as a percentage.

The controls also display economically plausible dispersion. Negative minima for ROA and Profit indicate that the sample includes loss-making firms rather than only successful policy recipients. Growth is particularly volatile, which is consistent with uneven commercialization and cyclical demand. Fixed assets average 21.01%

© 2026 The Author(s). Authors retain copyright in their articles.

of total assets, reflecting the equipment intensity of advanced manufacturing. Pairwise correlations provide preliminary direction but cannot substitute for the multivariate bounded-outcome model. A variance-inflation-factor check based on the baseline covariates and year indicators yields 1.02 for Tax, a maximum of 3.99 among firm-level covariates, and 4.91 on average when year indicators are included. All values remain below 10, so the estimates do not display severe multicollinearity under the conventional diagnostic threshold.

4.2 Baseline Estimates

Table 4 presents the two-limit Tobit estimates. Column (1), which includes only Tax, produces a coefficient of 0.4932 with a standard error of 0.1249. Column (2) adds the six firm controls and increases the coefficient to 0.5213. Column (3) additionally controls for year effects, producing a coefficient of 0.5549 with a standard error of 0.1114. All three coefficients are positive at the 1% level. The sign and significance are therefore not explained by the included differences in size, profitability, growth, fixed-asset intensity, board independence, or common annual shocks. The evidence supports H1 as a robust conditional association within the baseline specification.

Because the reported coefficient belongs to the latent Tobit equation, economic magnitude is evaluated using the average marginal effect on the expected observed score. The estimate is 0.5039. Since Tax is a ratio, a one-percentage-point increase in refunds relative to revenue is associated with an average increase of approximately 0.0050 in expected Eff, holding the other covariates and year effects constant. Relative to the sample mean of 0.2393, this change is not trivial, but it should not be read as a causal policy multiplier. It describes the average model-implied difference associated with the observed support intensity.

Table 4: Baseline two-limit Tobit estimates

Variable	(1) Eff	(2) Eff	(3) Eff
Tax	0.4932***	0.5213***	0.5549***
	(0.1249)	(0.1107)	(0.1114)
Size	-	-0.0498***	-0.0497***
		(0.0019)	(0.0019)
ROA	-	0.4945***	0.5268***
		(0.0803)	(0.0809)
Profit	-	0.1462***	0.1377***
		(0.0353)	(0.0355)
Growth	-	-0.0722***	-0.0769***
		(0.0062)	(0.0065)
Fix	-	-0.1908***	-0.1937***
		(0.0192)	(0.0192)
Indep	-	-0.0003	-0.0002
		(0.0004)	(0.0004)
Firm controls	No	Yes	Yes
Year effects	No	No	Yes
Firm-clustered SE	Yes	Yes	Yes
N	16,817	16,817	16,817

*Note: Firm-clustered robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

The controls clarify what the efficiency score captures. ROA and Profit are significantly positive, indicating that stronger asset use and profit conversion are associated with a more favorable relationship between debt inputs and financial outputs. Size is negative in the controlled models. Large firms can obtain more credit, but access alone may create organizational slack, project proliferation, or debt accumulation that is not matched by return. Growth is also negative. Rapid revenue expansion often requires working capital and capacity ahead of realized earnings, so high-growth firms may temporarily appear less efficient when debt rises before output measures adjust.

Fix is negative and significant. Tangible assets can support collateral, yet specialized equipment is costly, illiquid, and subject to long recovery periods. If equipment expansion is debt financed and return on invested capital responds slowly, the contemporaneous input-output relationship deteriorates. Indep is statistically insignificant, suggesting that the percentage of independent directors alone does not capture the quality of monitoring relevant to debt allocation. Overall, the controls reinforce a central interpretation: financing efficiency is not equivalent to financing volume. Debt contributes to efficiency only when cost, scale, maturity, and operating return remain aligned.

© 2026 The Author(s). Authors retain copyright in their articles.

4.3 Robustness and Mechanism

Table 5 separates checks that preserve the central measurement design from tests that alter it. When Tax is lagged by one period, the coefficient is 0.5272 with a standard error of 0.1074 and remains significant at 1%. The estimate is close to the baseline value of 0.5549. This result weakens, but does not eliminate, the concern that contemporaneously efficient firms simply receive larger refunds in the same year. It is also consistent with a mechanism that requires time: retained cash can affect refinancing, project continuity, and creditor assessment beyond the period in which the refund is received.

Reapplying the 1st- and 99th-percentile winsorization produces a Tax coefficient of 0.5524, again close to the baseline. This indicates that the main association is not driven by a small number of extreme observations. The exercise is best understood as a consistency check because the original estimation sample was already winsorized. It nevertheless confirms that the reported coefficient is reproducible under the stated treatment of outliers.

The two alternative definitions do not support the expected positive direction: the coefficient is -0.9910 when refunds are scaled by total assets and -0.0773 when Eff2 replaces the DEA score, with both estimates statistically significant. These reversals may arise because the total-assets denominator embeds capital intensity and Eff2 does not preserve the CCR input-output meaning; they are therefore treated as measurement sensitivity, while the lagged and re-winsorized estimates remain the principal robustness evidence supporting H1.

Table 5: Robustness and measurement-sensitivity checks

Specification	Outcome	Tax measure	Coefficient (SE)	N
Baseline full model	Eff	Tax	0.5549*** (0.1114)	16,817
Scaled by total assets	Eff	Tax2	-0.9910*** (0.1428)	16,817
Alternative efficiency measure	Eff2	Tax	-0.0773*** (0.0189)	16,817
One-period lag	Eff	L.Tax	0.5272*** (0.1074)	1.37 x 10^4
Re-winsorized sample	Eff	Tax	0.5524*** (0.1109)	16,817

*Note: Every specification includes firm controls and year effects. *** indicates significance at the 1% level.*

The mediation analysis examines whether financing-constraint relief is one pathway linking tax support to efficiency. Column (1) of Table 6 restates the total Tax coefficient of 0.5549. In column (2), Tax is positively related to ASA: the coefficient is 1.4739 with a standard error of 0.0884. Because larger ASA values denote weaker constraints, the result is consistent with tax refunds widening the feasible financing set. In column (3), ASA is positively associated with Eff at 0.0568, while the Tax coefficient remains positive but falls to 0.4713. The coefficient pattern is consistent with partial, rather than full, mediation.

The product of the Tax-to-ASA and ASA-to-Eff coefficients is 0.0837, which equals approximately 15.09% of the total Tax coefficient. This share is economically plausible because liquidity relief is only one channel. Tax support may also affect efficiency through refinancing timing, creditor signals, profitability, or project continuity, while some funds may not affect debt decisions at all. The result supports H2, subject to two qualifications. The available output does not provide a bootstrap confidence interval for the indirect effect, and the mediator may share unobserved determinants with the outcome. The mediation evidence is therefore mechanism-consistent rather than definitive causal decomposition.

Table 6: Financing-constraint mediation

Variable	(1) Eff	(2) ASA	(3) Eff
Tax	0.5549***	1.4739***	0.4713***
	(0.0556)	(0.0884)	(0.0558)
ASA	-	-	0.0568***
			(0.0048)
Firm controls	Yes	Yes	Yes
Year effects	Yes	Yes	Yes
N	16,817	16,817	16,817
Indirect effect (a x b)	-	-	0.0837
Share of total effect	-	-	15.09%

*Note: Standard errors are in parentheses. ASA is the absolute SA index, so a larger value denotes weaker financing constraints. *** indicates significance at the 1% level.*

© 2026 The Author(s). Authors retain copyright in their articles.

4.4 Heterogeneity

Average effects may conceal differences in external finance, implicit guarantees, and technological needs. Table 7 therefore reports subgroup coefficients and independent-sample coefficient-difference z-tests. The difference test is essential because statistical significance in one subgroup and insignificance in another does not by itself establish that the coefficients differ. Every subgroup model retains the firm controls and year effects used in the baseline.

Tax is positive in both the eastern and non-eastern samples, but their coefficient difference is not significant ($p = 0.498$); the same holds for smaller and larger firms ($p = 0.612$), so neither region nor firm size exhibits statistically supported heterogeneity. National policy transmission and revenue scaling may reduce cross-group differences, while the broad regional classification and median size split may conceal finer variation.

Ownership produces a statistically meaningful contrast. The coefficient for state-owned enterprises is -0.0840 and insignificant, whereas the coefficient for non-state-owned enterprises is 0.2600 and significant at 1%. The difference p-value is 0.004. Non-state-owned advanced manufacturers often have weaker implicit guarantees and face stricter collateral and information requirements. A tax refund can therefore have greater marginal value by reducing high-cost borrowing and supporting credit assessment. State-owned firms generally have broader financing access, so additional tax support may be less likely to alter their debt allocation at the margin.

Technology level produces the largest difference. The Tax coefficient is 0.4519 for high-technology firms but -0.0741 and insignificant for the lower-technology group; the p-value for the difference is below 0.001. High-technology firms have greater research intensity, more intangible assets, and longer intervals between expenditure and revenue. Timely tax relief can preserve technical programs and improve the continuity that creditors evaluate. For lower-technology firms, the use of additional funds may depend more on capacity demand and cost control, so the same refund measure does not generate a detectable efficiency association. These results identify ownership and technology, rather than region or scale, as the dimensions with statistically supported heterogeneity.

Table 7: Heterogeneity in the Tax coefficient

Dimension	Group 1: Tax (SE)	Group 2: Tax (SE)	N1 / N2	Difference	p-value
Region	East	Non-East	12,403 / 4,414	-0.0785	0.498
	0.1877*** (0.0607)	0.2662*** (0.0988)			
Firm size	Small	Large	8,409 / 8,408	0.0507	0.612
	0.1955*** (0.0539)	0.1448* (0.0841)			
Ownership	State-owned	Non-state-owned	3,734 / 13,083	-0.3440	0.004
	-0.0840 (0.1032)	0.2600*** (0.0594)			
Technology	High-tech	Lower-tech	9,224 / 7,593	0.5260	<0.001
	0.4519*** (0.0632)	-0.0741 (0.0806)			

*Note: p-values are from independent-sample coefficient-difference z-tests. All subgroup models include firm controls and year effects. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

5. Conclusion

Using 16,817 observations for Chinese A-share listed advanced manufacturing firms from 2015 to 2025, this study finds that tax refunds relative to operating revenue are positively associated with annual CCR debt-financing efficiency. The association remains positive after firm controls and year effects are included. The average marginal effect indicates that a one-percentage-point increase in Tax corresponds to an approximately 0.0050 increase in expected Eff. Financing-constraint relief accounts for an estimated 15.09% of the total coefficient. The association is significantly stronger for non-state-owned and high-technology firms, while regional and firm-size coefficient differences are not statistically significant.

The findings have several practical implications. Tax authorities can improve the financial value of support by making eligibility, calculation, and receipt more predictable, especially where firms must fund long research and equipment cycles. Policy evaluation should examine not only the amount of relief but also debt cost, maturity alignment, and capital return. Firms should integrate refunds into capital budgeting and prioritize projects with credible technical milestones and cash-flow paths rather than treating tax support as unrestricted

© 2026 The Author(s). Authors retain copyright in their articles.

surplus. Lenders may use stable and compliant tax-support records as supplementary information, but such records should complement rather than replace project-level cash-flow and repayment analysis.

The conclusions remain bounded by the research design. Tax refunds do not capture every form of income-tax reduction, and alternative scaling changes the result. CCR efficiency depends on selected inputs and outputs, the constant-returns assumption, and the annual comparison set. The two-stage Tobit model does not remove dependence induced by the estimated frontier, unobserved firm heterogeneity, or reverse causality. Future research should compare BCC and slack-based measures, build broader policy indicators, report bootstrap confidence intervals for mediation, implement Simar-Wilson inference, and exploit plausibly exogenous policy changes. Until those extensions are completed, the evidence should be read as a carefully qualified association that is informative for policy design but not a definitive causal estimate.

References

[1] Stiglitz J E, Weiss A. Credit rationing in markets with imperfect information. American Economic Review. 1981; 71(3):393-410.

[2] Myers S C, Majluf N S. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics. 1984; 13(2):187-221.

[3] Modigliani F, Miller M H. Corporate income taxes and the cost of capital: A correction. American Economic Review. 1963; 53(3):433-443.

[4] Zhai H, Yang M, Chan K C. Tax incentives, financing constraints and investment efficiency. Frontiers in Psychology. 2022; 13:899240.

[5] Occhino F. Firm tax cuts, financing frictions and aggregate implications. Economic Modelling. 2023; 121:106217.

[6] Hadlock C J, Pierce J R. New evidence on measuring financial constraints: Moving beyond the KZ index. Review of Financial Studies. 2010; 23(5):1909-1940.

[7] Charnes A, Cooper W W, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research. 1978; 2(6):429-444.

[8] He P, Liu Y. The impact of corporate ESG performance on debt-financing efficiency: Evidence from 963 listed companies. West China Finance. 2024; (9):3-10. [In Chinese].

[9] Tobin J. Estimation of relationships for limited dependent variables. Econometrica. 1958; 26(1):24-36.

[10] Simar L, Wilson P W. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics. 2007; 136(1):31-64.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.6166
Published by: Zeus Press

Executive Overseas Background, Customer Concentration, and Firm Internationalization: Evidence from Structural Changes Around 2018

Zikun Wu*

School of European and Latin American Studies, Shanghai International Studies University, Shanghai 201620, China

**Corresponding author: Zikun Wu.*

Abstract

Using panel data on Chinese A-share manufacturing firms from 2010 to 2023, this study examines whether executives' overseas background affects firm internationalization and whether customer concentration changes this relationship. The results show that executives' overseas background significantly increases the share of overseas revenue. Customer concentration strengthens this effect during 2012 and 2017, whereas the positive moderating effect weakens significantly and becomes statistically insignificant from 2018 onward. The findings are mainly driven by overseas work experience and remain robust to alternative fixed effects, an extended pre-2018 window, and IPWRA estimation. The study identifies customer dependence and changes in the external business environment as important boundaries of the value of international managerial human capital.

Keywords

executive overseas background, firm internationalization, customer concentration, resource dependence

1. Introduction

Enterprise internationalization is a process of accumulating overseas knowledge, investing resources, and gradually expanding business scope [1]. The ability of senior executives to identify overseas opportunities, understand institutional differences, and organize cross-cultural communication will influence a company's overseas decision-making. Existing research has found that executives' international experience is related to exports, cross-border mergers and acquisitions, operation of overseas subsidiaries and internationalization breadth [2–6], and the latest research has extended to investment destination selection [7].

Overseas experience for executives is not equally valuable in all companies. Customer concentration can bring stable orders and collaborative efficiency, but it can also make companies dependent on a few core customers and reduce market autonomy [8–10]. Existing research has noted the link between supply chain relationships and internationalization [11,12], but less has directly examined whether customer dependence alters the role of executives with overseas backgrounds and their temporal differences. This paper uses Chinese A-share listed manufacturing companies from 2010 to 2023 as a sample to examine the main effect, the moderating effect of customer concentration and its changes before and after 2018, thereby connecting the perspectives of high-level teams and resource dependence.

© 2026 The Author(s). Authors retain copyright in their articles.

2. Theoretical Analysis and Research Hypotheses

2.1 Overseas Background of Senior Executives and Firm Internationalization

The upper echelons theory suggests that managers' experience and cognition can influence a company's strategic judgment [13]. Overseas experience can increase understanding of foreign systems, market rules, and cultural differences, and help build overseas networks, international social networks, or international business connections. Early studies have found that international experience in management teams can accelerate overseas sales and expand the internationalization of enterprises [14]. Subsequent research has further demonstrated that executives' international experience can influence corporate internationalization through channels such as international information acquisition and strategic decision-making [15,16]. Research on Chinese companies has also found that overseas backgrounds of senior executives can promote corporate exports and overseas operations [17–19]. Based on this, it is proposed that:

H1: Overseas background of senior executives significantly improves the internationalization of enterprises.

2.2 The Moderating Effect of Customer Concentration

Resource dependence theory suggests that the lower the substitutability of external key resources, the stronger the power constraints on enterprises [20]. A high concentration of customers means that sales revenue depends more on a few customers, and the company has a stronger need to find new markets and channels. Therefore, the international knowledge, relationship networks and communication skills of executives with overseas backgrounds have greater compensatory value. Existing research has shown that the impact of customer and supplier concentration is constrained by the relationship between resource dependence and power balance [21], and digital transformation and overseas experience can also change supply chain configuration and corporate power [22,23]. Based on this, it is proposed that:

H2: During periods of relatively stable external operating environment, customer concentration positively moderates the relationship between the overseas background of senior executives and the degree of internationalization of enterprises.

2.3 Periodic Variations in Regulatory Effects

During the sample period, the international business environment and supply chain relationships continued to change, and the conditions for transforming existing overseas relationships and communication experience may change accordingly, as may the cost of expanding alternative customers [11,21]. The effects of overseas education and overseas work experience are not consistent [24]. Therefore, the moderating effect of customer concentration on the role of executives' overseas background may vary over time. Based on this, it is proposed that:

H3: The positive moderating effect of customer concentration weakens significantly in the later stages of the sample.

3. Research Design

3.1 Samples and Variables

This paper selects A-share listed manufacturing companies in China from 2010 to 2023, excludes companies with ST, *ST, abnormal operating status and missing key variables, and shrinks the continuous variables at the 1% and 99th percentiles. The degree of internationalization of an enterprise is the proportion of overseas operating revenue to total operating revenue; the overseas background of senior management is the proportion of senior management with overseas study or work experience; and the customer concentration is the proportion of sales from the top five customers to annual sales. Explanatory variables, moderating variables, and control variables are all lagged by one period. Control variables include size, liabilities, profitability, growth, cash flow, fixed assets, age at listing, ownership, equity, and board characteristics.

© 2026 The Author(s). Authors retain copyright in their articles.

3.2 Model Setting

The baseline model was used to examine the impact of executives' overseas background. The sample from 2012 to 2017 included an interaction term between executives' overseas background and customer concentration. Based on the structural change test of the candidate years, 2018 was the earliest node to show a significant structural difference, and the full sample was accordingly supplemented with Post2018 and a triple interaction term. The model controls for firm fixed effects and industry-year interaction fixed effects, and clusters firms by standard error.

4. Empirical Results

4.1 Descriptive Statistics and Benchmark Results

Table 1 shows that the mean internationalization level of enterprises is 0.0836, and the median is 0; the mean proportion of senior executives with overseas backgrounds is 0.0612; and the mean customer concentration is 0.3154. No severe collinearity was found in the correlation and VIF tests, and the average VIF was 1.409.

Table 1: Definitions of Main Variables and Descriptive Statistics

Variable	Definition	N	Mean	Std. Dev.
Degree of Firm Internationalization	Overseas operating revenue / Total operating revenue	20,389	0.0836	0.1814
Overseas Background of Senior Executives	Proportion of senior executives with overseas study or work experience	20,389	0.0612	0.1260
Customer Concentration	Sales of top five customers / Annual sales	20,389	0.3154	0.2058
Firm Size	Natural logarithm of total assets	20,389	22.0465	1.1510
Leverage	Total liabilities / Total assets	20,389	0.3947	0.1881
Return on Assets (ROA)	Net profit / Total assets	20,389	0.0463	0.0608

Note: The main explanatory variables, moderating variables, and control variables all use a one-period lag; other control variables are detailed in the main text.

In Model (1) in Table 2, the coefficient for overseas background of senior executives is 0.0648, which is significant at the 1% level, supporting H1. This indicates that when the proportion of overseas background among the senior executives of the same company increases, the proportion of overseas operating revenue also increases.

4.2 Moderating Effect and Time Variation

In model (2), the interaction term between executives' overseas background and customer concentration is 0.5681, which is significant at the 5% level, and H2 is supported. When the customer concentration is at the 75th percentile, the marginal effect of executives' overseas background is 0.1287, and it is significant at the 5% level; it is not significant at lower concentration levels, indicating that the value of their resources is mainly reflected when customer dependence is strong.

The triple interaction term of model (3) is -0.3935, which is significant at the 5% level. The linear combination test showed that the moderating effect before 2018 was 0.3214, which was significant at the 5% level; the net moderating effect from 2018 to 2023 was -0.0721, but not significant. Therefore, the original positive moderating effect was significantly weakened and lost statistical significance, rather than turning into a significant negative effect, which supports H3.

Table 2: Core Regression and Robustness Tests

Variable	(1) Baseline Regression	(2) 2012–2017	(3) Structural Change	(4) Robustness
Overseas Background of Senior Executives	0.0648*** (0.0217)	0.0715 (0.0497)	0.0165 (0.0284)	Overseas work background 0.0837***(0.0269)
Customer Concentration	—	-0.0038 (0.0277)	-0.0356* (0.0196)	—

© 2026 The Author(s). Authors retain copyright in their articles.

Variable	(1) Baseline Regression	(2) 2012–2017	(3) Structural Change	(4) Robustness
Overseas Background × Customer Concentration	—	0.5681** (0.2445)	0.3214** (0.1575)	—
Overseas Background × Post2018	—	—	0.0657** (0.0293)	—
Customer Concentration × Post2018	—	—	0.0418** (0.0179)	—
Triple Interaction Term	—	—	-0.3935** (0.1812)	—
IPWRA Treatment Effect	—	—	—	0.0081*** (0.0019)
Control Variables	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	—
Industry × Year Fixed Effects	Yes	Yes	Yes	Industry, Year
Number of Observations	21,041	7,092	20,389	20,699

*Note: Robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Model (4) reports the alternative variable of overseas work background and the IPWRA treatment effect.*

4.3 Robustness and Outcome Boundaries

The coefficient for the proportion of senior executives with overseas work experience is 0.0837, which is significant at the 1% level, while the proportion of those with overseas education experience is not significant. After replacing the fixed effects and expanding the initial sample window, the core results remain unchanged. IPWRA estimates show that companies with executives from overseas backgrounds have an average increase in internationalization of 0.0081%, which is significant at the 1% level. After changing the dependent variable to whether overseas business is conducted or the logarithm of overseas operating revenue, the result was not significant. Supplier concentration and overall supply chain concentration did not exhibit the same moderating effect. The conclusions of this paper are mainly applicable to the degree of internationalization reflected by the proportion of overseas revenue, and the contextual effect is concentrated on customer dependence.

5. Conclusions and Implications

Research has found that overseas backgrounds of senior executives significantly promote the internationalization of manufacturing companies. Between 2012 and 2017, customer concentration amplified this effect, but after 2018, this positive moderating effect weakened significantly. Overseas experience is not a fixed resource that can function independently of the business environment. Its value depends on market dependence and external operating conditions.

Companies with a high concentration of customers can give priority to hiring senior executives with overseas work experience and international market networks. However, with increased volatility in the external environment, companies also need to build organizational-level capabilities in overseas channels, digital marketing, and supply chain early warning. This paper identifies overseas backgrounds based on publicly available resumes, making it difficult to measure the country of experience, duration, and quality. Changes around 2018 may be related to multiple external business environment factors simultaneously, and this paper cannot identify their specific sources, therefore it cannot be explained as a strict causal effect of a single shock.

References

[1] Johanson, J., & Vahlne, J.-E. (1977). The internationalization process of the firm: A model of knowledge development and increasing foreign market commitments. Journal of International Business Studies, 8(1), 23–32. https://doi.org/10.1057/palgrave.jibs.8490676.

[2] Ren, G., Chen, W. H., & Zhong, X. (2019). Executive international experience, environmental uncertainty and the process of enterprise internationalization. Foreign Economics & Management, 41(9), 109–121. https://doi.org/10.16538/j.cnki.fem.20190808.005.

© 2026 The Author(s). Authors retain copyright in their articles.

[3] Li, Q., Wang, J. L., & Yang, J. (2024). CEO overseas background and firm exports. International Business Research, (6), 13–24.

[4] Ye, J. S., Pang, S. P., & Zhong, C. B. (2024). Executive overseas background and corporate cross-border M&A behavior: Empirical evidence from China. World Economy Studies, (7), 91–107. https://doi.org/10.13516/j.cnki.wes.2024.07.004.

[5] Agcayazi, B., Hibbert, A. M., & Morillon, T. G. (2024). CEO international background and cross-border M&As. Journal of Banking & Finance, 164, 107129. https://doi.org/10.1016/j.jbankfin.2024.107129.

[6] Sun, X., & Sun, H. (2025). Can CEOs' overseas background facilitate enterprise internationalization breadth? Finance Research Letters, 86, 108397. https://doi.org/10.1016/j.frl.2025.108397.

[7] Coda Zabetta, M., Miguelez, E., Lissoni, F., et al. (2025). Foreign-origin managers and FDI location choice. Journal of International Business Studies, 56(2), 194–202. https://doi.org/10.1057/s41267-024-00726-2.

[8] Patatoukas, P. N. (2012). Customer-base concentration: Implications for firm performance and capital markets. The Accounting Review, 87(2), 363–392. https://doi.org/10.2308/accr-10198.

[9] Irvine, P. J., Park, S. S., & Yildizhan, C. (2016). Customer-base concentration, profitability, and the relationship life cycle. The Accounting Review, 91(3), 883–906. https://doi.org/10.2308/accr-51246.

[10] Hua, Z. B., Ren, S. C., & Liu, Y. D. (2026). Customer concentration: Performance characteristics, formation logic and research prospects. East China Economic Management, 40(7), 120–128. https://doi.org/10.19629/j.cnki.34-1014/f.250828035.

[11] Mao, J., & Yang, S. (2024). Changes in supply chain relationships and the enterprise internationalization process. Research in International Business and Finance, 67, 102148. https://doi.org/10.1016/j.ribaf.2023.102148.

[12] Wang, C., & Feng, C. (2025). Digital technology and degree of internationalization—The mediating role of supply chain concentration. Finance Research Letters, 85, 107875. https://doi.org/10.1016/j.frl.2025.107875.

[13] Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. Academy of Management Review, 9(2), 193–206. https://doi.org/10.5465/amr.1984.4277628.

[14] Reuber, A. R., & Fischer, E. (1997). The influence of the management team's international experience on the internationalization behaviors of SMEs. Journal of International Business Studies, 28(4), 807–825. https://doi.org/10.1057/palgrave.jibs.8490120.

[15] Pisani, N., Muller, A., & Bogăţan, P. (2018). Top management team internationalization and firm-level internationalization: The moderating effects of home-region institutional diversity and firm global focus. Journal of International Management, 24(3), 239–256. https://doi.org/10.1016/j.intman.2018.01.002.

[16] Azam, A., Bertolotti, F., Boari, C., et al. (2020). Top management team international experience, international information acquisition and international strategic decision rationality. Review of International Business and Strategy, 30(3), 441–456. https://doi.org/10.1108/RIBS-01-2020-0010.

[17] Qi, J. H., & Yang, W. H. (2020). The impact of overseas background directors on firm export behavior: Incremental quantity or quality improvement? Finance and Trade Research, (7), 34–49. https://doi.org/10.19337/j.cnki.34-1093/f.2020.07.003.

[18] Yao, K., & Wang, Y. J. (2020). Returnee executives and firm internationalization: An empirical study based on China's high-tech listed companies. Economic Theory and Business Management, (11), 55–71.

[19] Yi, C. J., Zhao, X. Y., & Huang, Y. (2022). Executive overseas background, entry mode and the survival performance of multinational enterprises' overseas subsidiaries. Economic Theory and Business Management, (8), 88–100.

[20] Pfeffer, J., & Salancik, G. R. (1978). The External Control of Organizations: A Resource Dependence Perspective. Harper & Row.

© 2026 The Author(s). Authors retain copyright in their articles.

[21] Jiang, S., Yeung, A. C. L., Han, Z., et al. (2023). The effect of customer and supplier concentrations on firm resilience during the COVID-19 pandemic: Resource dependence and power balancing. Journal of Operations Management, 69(3), 497–518. https://doi.org/10.1002/joom.1236.

[22] Wu, Q., & Yao, Y. X. (2023). Enterprise digital transformation and supply chain configuration: Concentration or diversification. China Industrial Economics, (8), 99–117. https://doi.org/10.19581/j.cnki.ciejournal.2023.08.005.

[23] Li, C., Shi, E. Y., & Wang, X. (2025). Research on the mechanism of executive overseas experience affecting firms' supply chain discourse power. Collected Essays on Finance and Economics, (3), 112–124. https://doi.org/10.13762/j.cnki.cjlc.20241126.002.

[24] Li, X. Q., & Ni, M. M. (2025). The impact of executive overseas background on firm internationalization performance: The moderating role of incentive mechanisms. Journal of Chongqing Technology and Business University (Social Sciences Edition), (1), 168–180. https://doi.org/10.3969/j.issn.1672-0598.2025.01.015.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.6773
Published by: Zeus Press

The Spatial and Temporal Distribution Characteristics and Governance Tension of Tricycle Passenger Transport in County Towns from the Perspective of Informal Economy—A Case Study of A Certain County in Guangdong Province

Yue Li*

School of Digital Economics, Guangzhou Xinhua University, Guangzhou, 510520, China

**Corresponding author: Yue Li.*

Abstract

In the process of county-level urbanization, tricycle passenger transport, as a form of informal economy, effectively fills the spatial and temporal gaps in public transportation. This study selects a county in Guangdong Province as a case, and comprehensively uses spatial analysis and semi-structured interviews to examine the spatial and temporal evolution characteristics and governance contradictions of tricycle passenger transport. The results show that the industry presents a 'high-density, small-scale' 'ring-locking' spatial form, which is highly concentrated in the central area of the county in terms of spatial distribution. Although some practitioners hold business licenses and meet the legal requirements in form, they still encounter institutional exclusion and barriers to career advancement in actual operation, and are in the marginalized position of "legal but excluded." This paper argues that the current 'de-informalization' strategy ignores the social security supplementary function undertaken by the industry. In the future, we should realize the transformation from "compulsory banning" to "orderly counseling," recognize its existence value, implement a flexible management mechanism, and separate social security from formal employment system, so as to promote the inclusive transformation of governance model.

Keywords

informal economy, tricycle passenger transport, spatial distribution, governance tensions, case interviews

1. Introduction

As county towns rapidly undergo urbanization, the formal public transportation system—due to an imbalance between costs and benefits—is struggling to cover all spatial and temporal blind spots. In this gap, tricycle passenger transport has emerged as an economic activity that requires low barriers to entry, offers great flexibility, and does not necessitate complicated approvals, swiftly filling the market void. From the perspective of the informal economy, tricycle operators are typically self-employed workers [1]. They use their own tricycles as a production tool, participate in market transactions through self-employment, and carry out business activities in the absence of formal employment contracts and social security systems to support family livelihoods.

© 2026 The Author(s). Authors retain copyright in their articles.

However, the current institutional arrangements are often ambiguous about informal employment [2], and even regard it as a “stain” in the process of urbanization, trying to remove it by administrative means. Such governance approaches ignore the important role of the informal economy in absorbing low-skilled labor and reducing urban operating costs. Some studies have pointed out that the proportion of urban informal employment in total employment in China is between 33.2 % and 44.7 %; in view of its huge volume, the ‘de-informalization’ strategy may not be effective in practice [3]. The existing literature pays more attention to the violation or safety risk of tricycle passenger transport [4], but few questions from the perspective of economic rationality: why does the informal economy continue to exist under strict control? How does its spatial distribution reflect the practitioners’ continuous pursuit of economic efficiency?

Based on this, this study takes a typical county in Guangdong Province (highly concentrated tricycle passenger transport activities) as a case. The tricycle passenger transport economy in the county presents the distribution characteristics of ‘local high density’ in space. The purpose of the study is to explore: as a typical form of informal economy, how does tricycle passenger transport respond to market risks and administrative constraints with spatial strategies? Has the regulatory behavior represented by franchising changed the inherent vulnerability of its informal attributes?

2. Theory Framework

The theory of the informal economy holds that the existence of the informal economy is not due to backwardness, but rather represents a structural feature of the capitalist development process [5]. For the operators of tricycle passenger transport in the county town, their economic behavior follows the following logic:

2.1 Low Barriers to Entry and Excessive Competition

Due to the lack of technical barriers and capital thresholds, a large influx of surplus labor has led to "involutionary" competition within the tricycle passenger transport industry [6].

2.2 Flexibility and Vulnerability Coexist.

Although working hours are flexible, a lack of bargaining power makes operators of tricycle passenger transport highly vulnerable to external shocks, leading to high volatility in the number of practitioners and fragile livelihoods.

2.3 Spatial Dependency

The informal economy cannot, like formal enterprises, have a fixed business location; instead, it must rely on public spaces with high foot traffic to carry out its mobile operations. The survival of tricycle passenger transport depends heavily on the occupation and utilization of urban space.

3. Research Methodology

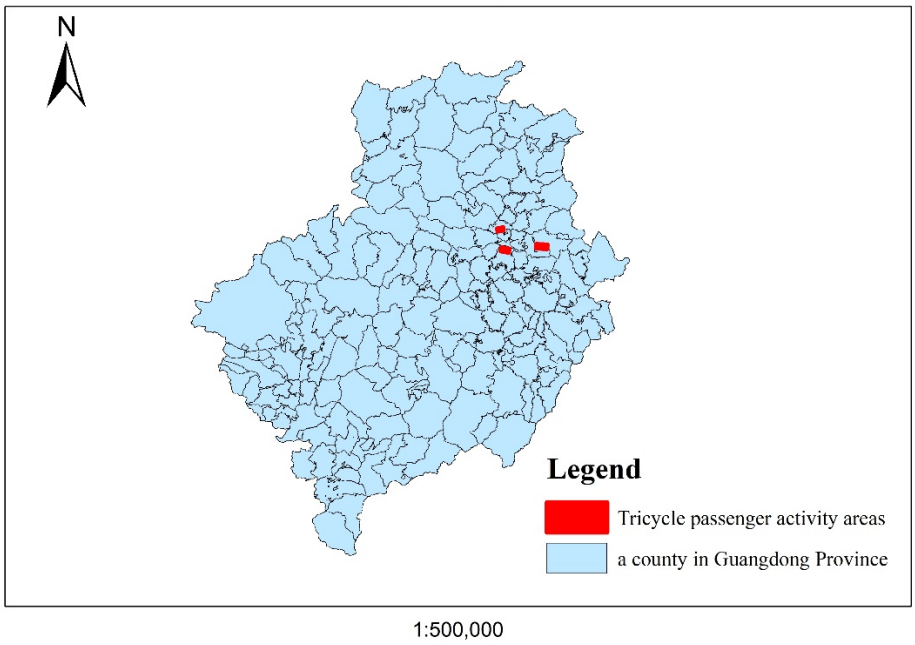
This study comprehensively uses the mixed method design of spatial analysis and case interview to achieve the mutual verification between macro spatial pattern and micro individual behavior.

3.1 Spatial Analysis Method

As the core method of analyzing the spatial distribution of social and economic phenomena in the field of geography and urban planning, spatial analysis is often used to reveal the location characteristics. With the help of ArcGIS platform, based on the vector data of Tianditu 2024 and the administrative boundary data of Guangdong Province, this study implements spatial visualization of three-wheel passenger transport activities in a county. By using the ‘point to polygon’ conversion technology, the scattered passenger transport behavior is transformed into the geographic information layer, and then the spatial aggregation situation is identified.

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 1: A county tricycle passenger transport area distribution map



It can be seen from Figure 1 that the tricycle passenger transport activities are densely distributed in three adjacent red areas in the center of the county, and its area is less than 2 % of the total area of the county's administrative division. This area covers the county people's hospital, the central farmers' market and the commercial square, which is not only the most densely populated area in the county, but also the most lacking area of public transport service coverage. This distribution situation shows that tricycles are frequently used for passenger transport; this spatial dispersion is not randomly scattered, but shows a structural 'spatial locking' pattern.

3.2 Case Interview Method

Case interview method is a qualitative research method commonly used in social science. With the help of in-depth interactive communication, it can grasp the subject meaning cognition and action logic of the research object. In view of the fact that statistical data are difficult to capture the actual life experience of individuals in the institutional cracks due to their macro attributes, this study selected a typical representative tricycle operator to conduct semi-structured in-depth interviews.

3.2.1 SAMPLE Selection and Representativeness

The respondent is a 52-year-old male with a junior high school education who has been engaged in passenger transportation operations since 1996 and holds a "Non-Motorized Vehicle Operation Permit" issued by the county government. This sample was selected based on the following considerations:

(1) Long-term experience: With 30 years of professional experience, he has fully witnessed the entire process of passenger transport by three-wheeled vehicles in the county town—from its wild, unregulated growth to its strict regulation and control—a period spanning a considerable length of time.

(2) Compliance: Although these practitioners are officially certified and legally authorized, they are at the boundary between formal systems and informal practices. Its situation of "holding certificates but being marginalized" is precisely the most direct mapping of the current governance tension.

(3) Typicality: its occupational substitution space is extremely narrow (such as 'can only do security'), which directly reflects the structural dilemma faced by low-skilled labor groups in county-level cities [7].

© 2026 The Author(s). Authors retain copyright in their articles.

3.2.2 Interview Implementation and Analysis

The interview is carried out layer by layer around the five dimensions of "employment motivation," "time and space trajectory," "income level," "administrative contact" and "future expectation." This study implements grounded theory coding for interview records, aiming to extract the 'livelihood resilience' and 'identity cognition' of informal employees from the interviewees' oral experience, so as to make up for the limitations of spatial analysis methods in explaining individual behavior motivation, and ultimately serve the research goal of human-land relationship coordination.

4. Spatial and Temporal Distribution Characteristics and Their Logic of Passenger Transport by Tricycles in County Towns

4.1 Spatial Agglomeration: The "Core Circle" That is Locked In

The map clearly shows that the passenger-carrying activities of three-wheeled vehicles are not evenly distributed across the entire county; rather, they are highly concentrated in the core area of the county seat. This "circle lock-in" pattern is not the result of the operator's active planning, but the product of the interaction of multiple forces. First, because tricycle passenger transport is billed by mileage and the pricing is low, the operating income is highly dependent on the passenger flow, so the vehicle naturally gathers to the crowded nodes such as hospitals, markets, and transportation hubs. Secondly, the urban management department imposes tricycle passenger transport activities on a specific range by setting 'no-go zones', 'strict road control' and other control methods, thus forming a spatial governance logic of 'allowing its presence and limiting its diffusion'.

From the perspective of human-earth system coupling, this kind of spatial locking represents a non-equilibrium coupling state between human activities and urban spatial environment. Liu Yansui believes that contemporary human activities have had a significant impact on the human-land system on the earth's surface [8], and the coupling level within the human-land system is directly related to the sustainability of regional development. The over-reliance of tricycle passenger transport on the core area is, in essence, a form of "over-coupling"—human activities are confined to an extremely small spatial area, leading to strained human-land relationships and intensified competition.

The respondents' experiences confirm this spatial logic: "Previously, we could still venture a bit farther; now, as soon as we step beyond these few intersections, we're either turned back or even fined." This spatial contraction not only reduces its operating radius, but also blocks their way to achieve capital accumulation through business expansion, leaving them in a vicious cycle of 'short distance-low frequency-low income' for a long time.

4.2 Temporal Regularity: The Rhythm of Survival in the Gaps

In addition to the characteristics of spatial agglomeration, the working hours of tricycle passenger transport operators also show obvious characteristics of 'gap filling'. They often go out at 5 to 7 o'clock in the morning, 5 to 8 o'clock in the evening, and rainy days or holidays when public transport capacity is insufficient, so as to make up for the gap of conventional bus services. Such timing not only reflects market rationality, but also reflects the passive adjustment of regulatory pressure: avoiding the peak of law enforcement inspections and using regulatory space to protect its living space.

However, this 'survival in the cracks' mode of operation further aggravates its informal attributes. Practitioners of tricycle passenger transport can neither build a solid customer network nor accumulate industry reputation. Their income is completely dependent on the immediate flow of people in the past, and the unstable state of 'living today and not knowing' has gradually become the norm.

© 2026 The Author(s). Authors retain copyright in their articles.

5. The Marginalization of Compliers

5.1 The Paradox of Formal Recognition and Practical Exclusion

Holding a license to operate was supposed to be the system's formal recognition of rickshaw drivers' right to livelihood. Yet in practice, these licenses have failed to grant them equal road rights or secure operating conditions. On the contrary, they have become a kind of "limited permit"—allowing you to operate only at designated times, in designated locations, and in designated ways; once you step beyond those boundaries, you'll face penalties. This governance logic— "recognizing capability but not recognizing rights"—leaves licensed operators stuck in the awkward position of being "legal yet illegitimate."

Although the respondent has held the relevant documentation for 30 years, he has never been integrated into the social security system nor received any vocational training or support for career transitions. His "compliance" merely means "not being banned," rather than "being accepted." This disconnect between form and substance is precisely a manifestation of the deep-seated tensions inherent in the governance of the informal economy.

5.2 Occupational Lock-in and the Dilemma of Job Mobility

30 Years of work experience have made the respondents' human capital highly specialized in tricycle passenger transport, leaving them unable to translate this expertise into employability in other industries. At the same time, their income is just sufficient to cover basic living expenses, leaving them unable to afford the costs of skills training or starting their own businesses [9]. When asked about his future plans, he said, "I'll keep going as long as I can still do the job," with alternative options limited solely to security guard positions that offer similarly low pay and poor benefits.

This kind of career lock-in [10] is not merely a limitation of individual capabilities; rather, it's a product of the institutional environment. Informal employment lacks career advancement pathways, social security systems, and retraining mechanisms, making it extremely difficult for workers to exit once they've entered this sector. Meanwhile, urban governance's ongoing crackdown on tricycle passenger transport further narrows their room for transformation, creating a vicious cycle in which "the more they're regulated, the more trapped they become; the more trapped they become, the more dependent they grow."

6. Conclusion and Discussion

Through the combination of spatial analysis method and case interview method, this paper reveals the dual dilemma of tricycle carrying passengers in the county under the framework of informal economy : On the one hand, their activities are strictly confined to the narrow space of the county town's core area, resulting in "circular lock-in"; on the other hand, even after obtaining operational qualifications, practitioners still encounter institutional exclusion and barriers to occupational mobility, becoming trapped in a state of "compliance yet marginalization."

The issue of tricycle passenger transport is essentially a conflict between the goals of urban modernization and the survival logic of those at the grassroots level. Current governance places too much emphasis on "de-informalization.", relevant departments by driving them away, we purify the space while neglecting the safety-net function that the informal economy plays in county-level societies.

As far as the informal economy is concerned, the idea of follow-up governance should be shifted from 'containment' to 'diversion'.

6.1 Recognizing the Legitimacy of the Informal Economy

Tricycle passenger transport should not be regarded as the 'focus' of the city; on the contrary, it should be seen as a useful complement to public transport and a reservoir of jobs.

© 2026 The Author(s). Authors retain copyright in their articles.

6.2 Spatial Justice and Flexible Management

In non-central areas or designated time periods, 'tolerance zones' are set up for the informal economy, allowing tricycle passenger transport to operate independently within the framework of established rules, rather than being completely driven to desperate situations.

6.3 Decoupling of Social Security

Try to separate social security benefits from employment status, and provide flexible employees with reasonable medical and old-age security to reduce their livelihood risks, rather than rigidly classifying them into a non-existent 'formal employment' framework.

References

[1] Xue, Z. Y. (2000). Developing the informal sector and informal employment in China. *China Employment* (10), 4.

[2] Ding, J. H., Leng, X. L., Song, X. K., Hammer, B., & Xu, Y. H. (2001). The transplantation and development of the concept of informal employment in China. *Chinese Journal of Population Scienc*e, (6), 8–15.

[3] Chen, M. X., Huang, X. R., Huang, G. Z., & Yang, Y. S. (2021). New urbanization and informal employment: Scale, pattern, and social integration. *Progress in Geography*, 40(1), 50–60.

[4] Qian, X. F. (2020). Local governance strategies and the rise and fall of the informal economy--A case of the government of unlicensed passenger tricycles in city A (Master's thesis, Central China Normal University).

[5] Portes, Alejandro; Haller, William. The informal economy. *The Handbook of Economic Sociology*. 403-425,727. Princeton: Princeton University Press, 2005.

[6] Huang, Z. Z. (2009). China's neglected informal economy: Reality and theory. *Open Times*, (2), 23.

[7] Jiao, X. Y. (2015). An analysis of the "semi-urbanization" problem and countermeasures in the process of urbanization. *Contemporary Economic Management*, 37(3), 64–67.

[8] Liu, Y. S. (2020). Modern human-earth relationship and human-earth system science. *Scientia Geographica Sinic*a, 40(8), 1221–1234.

[9] Yuan, Z. H. (2003). An analysis of the urban vulnerable groups in underdeveloped areas. *Theoretical Study*, (6), 55–56.

[10] Li, Q. (1999). Occupational mobility of urban migrant workers in mainland China. *Sociological Studies*, (3), 9.

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.

© 2026 The Author(s). Authors retain copyright in their articles.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.7483
Published by: Zeus Press

Can Annual Report Tone Signal Audit Risk? Evidence from Large Language Model Textual Analysis of Chinese Listed Firms

Wen Guan, Yuxuan Zhou*, Jing Zhang, Junyi Zang

Nanjing Audit University, Nanjing 211815, China

**Corresponding author: Yuxuan Zhou.*

Abstract

Using a sample of Chinese A-share listed firms from 2021 to 2024, we employ a large language model (LLM) to measure the tone of approximately 20,000 Management Discussion and Analysis (MD&A) sections and examine whether narrative tone signals future audit risk. We document four findings. First, MD&A tone significantly and negatively predicts the probability of receiving a modified audit opinion in the following year; the effect survives controls for the current-year opinion, firm fundamentals, and industry and year fixed effects, and holds within the subsample of firms with clean current-year opinions. Second, the predictive power operates through forward-looking information embedded in tone: firms with more negative tone exhibit significantly lower profitability and a higher likelihood of losses in the subsequent year. Third, in a horse race between the LLM-based measure and a conventional dictionary-based measure built on an established Chinese financial sentiment lexicon, only the LLM tone retains predictive power, indicating that the audit-relevant component of tone resides in contextual meaning that word counts cannot capture. Fourth, an anonymization test addresses look-ahead bias: after stripping all firm identifiers from the text, re-scored tone remains highly correlated with the original measure (Pearson correlation 0.79) and preserves its predictive power, ruling out the concern that the model "remembers" firm outcomes from its training data. In contrast, tone explains neither audit fees nor audit report lag, suggesting that soft information shapes auditors' risk judgments rather than audit pricing. Our study provides a validation framework for LLM-based textual measures in accounting research and offers practical implications for using narrative disclosure in risk assessment.

Keywords

annual report tone, MD&A, audit opinion, large language models, textual analysis

1. Introduction

Beyond the financial statements, corporate annual reports contain extensive narrative disclosure. The Management Discussion and Analysis (MD&A) section, in which managers describe operations, risks, and prospects in their own words, is widely viewed as carrying incremental information that quantitative disclosures cannot fully convey [1-4]. How managers say things may matter as much as what they say: whether difficulties are downplayed or candidly acknowledged, and whether the outlook is stated with confidence or hedged with qualifications, may reveal managers' private expectations about the firm [5,6].

© 2026 The Author(s). Authors retain copyright in their articles.

Auditors are among the most important professional users of annual report information. The formation of an audit opinion is fundamentally a risk-assessment process: the auditor must evaluate the risk of material misstatement and the entity's ability to continue as a going concern. If narrative tone indeed contains forward-looking information about a firm's condition, firms with more negative tone should experience more adverse developments in subsequent periods and should therefore be more likely to receive modified audit opinions. Testing this conjecture, however, has long been constrained by measurement: tone is soft information, and the standard measurement technology, counting words against a sentiment dictionary, cannot process context, negation, qualification, or euphemism [8]. The resulting measurement error may obscure true relations, particularly in Chinese, where disclosure language is formulaic and indirect.

Large language models offer a new measurement technology. Unlike bag-of-words approaches, an LLM reads the full text and evaluates meaning in context, in principle approximating the judgment of a skilled human reader [9,10]. We use the DeepSeek LLM to score the tone of roughly 20,000 MD&A sections of Chinese A-share listed firms over 2021-2024 and ask three questions. First, does MD&A tone predict future modified audit opinions? Second, what mechanism underlies any predictive power? Third, does the LLM measure add value relative to the conventional dictionary approach?

We find that MD&A tone negatively and robustly predicts next-year modified audit opinions. A one-unit increase in tone (about three standard deviations) is associated with a 1.2-percentage-point decline in the probability of a modified opinion, against an unconditional rate of 3.3%. The effect is present, and indeed stronger, among firms whose current-year opinion is clean, so it is not driven by the persistence of opinions among already-troubled firms. Consistent with a forward-looking-information mechanism, tone predicts next-year return on assets and the incidence of losses after controlling for current profitability. In a horse race, the dictionary-based tone measure has no predictive power for future opinions, whereas the LLM measure remains significant with both measures included; the audit-relevant component of tone is precisely the part that word counting cannot extract. An anonymization experiment that rescores 2,000 MD&As after removing all firm identifiers yields tone scores highly correlated with the originals that retain predictive power, addressing the look-ahead-bias concern specific to LLM-based research. Finally, tone explains neither audit fees nor audit report lag in our sample, a contrast suggesting that soft information enters auditors' risk judgments but not audit pricing in the Chinese market, where fees are sticky and dominated by client size.

This paper contributes to three strands of literature. First, we contribute to the emerging literature applying LLMs to accounting and finance texts [9,10,20] by proposing and implementing a two-part validation framework (a horse race against an established dictionary and an anonymization test for look-ahead bias) that we believe should become standard practice for LLM-based measures constructed over samples that may overlap with model training data. Second, we extend the literature on the economic consequences of disclosure tone [2,4,6] to auditor decisions in an emerging market, documenting that tone predicts audit opinions but not audit fees, a dissociation between risk judgment and pricing that prior single-outcome studies could not reveal. Third, our results are relevant for regulators and practitioners: narrative tone is a machine-readable leading indicator of audit risk that can be incorporated into supervisory screening at negligible cost.

The remainder of the paper proceeds as follows. Section 2 reviews the literature and develops hypotheses. Section 3 describes the sample, tone measurement, and research design. Section 4 presents the main results, mechanism tests, the horse race, and validity checks. Section 5 reports robustness tests. Section 6 concludes.

2. Literature Review and Hypothesis Development

2.1 The Information Content of Disclosure Tone

Loughran and McDonald [2] develop finance-specific English word lists and show that 10-K tone relates to returns and fundamentals, initiating a large literature on disclosure sentiment (see [8] for a survey). Li [1] documents that annual report readability relates to earnings persistence, Feldman et al. [11] find that changes in MD&A tone convey information incremental to accruals and earnings surprises, Davis et al. [3] show that optimistic earnings-press-release language predicts future performance, and Li [5] shows that forward-looking statements in filings carry information about future firm outcomes. For Chinese text, Yao et al. [4] construct a finance-domain sentiment lexicon via dictionary reorganization and deep learning, and show that annual report tone predicts stock returns and crash risk. Xie and Lin [12] find that managerial tone in earnings

© 2026 The Author(s). Authors retain copyright in their articles.

communication conferences predicts Chinese firms' next-year performance, and Zeng et al. [13] show that annual report tone predicts subsequent insider trading, evidence that tone reflects managers' private information, though it can also be strategically managed for impression purposes [6].

2.2 Tone and Auditing

Auditing standards require the auditor to assess the risk of material misstatement and to evaluate going-concern uncertainty; auditor effort, pricing, and reporting respond to client risk [7,14]. Narrative disclosure is a key input into the auditor's understanding of the entity. U.S.-based studies find that negative or uncertain MD&A language is associated with higher audit fees and longer audit delay [15], consistent with tone proxying for inherent risk that triggers greater audit effort and risk premia. In China's institutional setting, however, audit pricing is heavily influenced by fee stickiness and market competition [16], so fees may respond weakly to soft information. The audit opinion, by contrast, directly reflects the auditor's professional risk judgment: modified opinions in China are informative to the market and responsive to client-level risk factors such as earnings management incentives and regulatory scrutiny [17-19], making the opinion a sharper outcome for detecting whether auditors impound narrative information.

2.3 LLMs and Textual Measurement

The fundamental limitation of dictionary methods is the bag-of-words assumption: "risks have been effectively controlled" and "facing major risks" receive similar scores [8]. LLMs acquire contextual language understanding through pretraining and can assess the overall tenor of a document. Recent studies report that LLM-based measures of corporate disclosure outperform dictionary and other traditional approaches: sentiment from GPT-3.5-generated summaries of filings explains market reactions better than sentiment from the original text [9], ChatGPT-inferred news sentiment predicts stock returns [10], and GPT-based earnings predictions from anonymized financial statements rival trained human analysts [20]. LLM-based measurement, however, faces a distinctive identification threat, namely look-ahead bias: the model's training corpus may postdate the sample period, so scores for recognizable firms may embed the model's "memory" of subsequent outcomes rather than a genuine reading of the text [21,22]. Any credible design must address this concern directly.

2.4 Hypotheses

The preceding discussion yields three testable hypotheses:

H1: More negative MD&A tone predicts a higher probability of receiving a modified audit opinion in the following year.

H2: The predictive power of tone derives from forward-looking information about fundamentals: firms with more negative tone experience worse future performance.

H3: LLM-measured tone has incremental predictive power relative to dictionary-based tone.

3. Research Design

3.1 Sample and Data

Our initial sample comprises all A-share listed firms over 2021-2024. Annual report full texts (approximately 20,500 documents) are obtained from public disclosure channels (cninfo.com.cn); audit fees, audit opinions, auditor identities, financial data, ultimate-controller type, and Shenwan industry classifications are from the Wind database. We programmatically extract the MD&A section from each report (accommodating historical formats such as "Discussion and Analysis of Operations" and "Report of the Board of Directors"), achieving a 98.9% extraction rate and 20,287 MD&A texts with a mean length of about 59,000 Chinese characters.

Following convention, we exclude financial firms and observations with missing tone or key variables, yielding 18,479 firm-year observations. Continuous variables are winsorized at the 1st and 99th percentiles.

© 2026 The Author(s). Authors retain copyright in their articles.

Tests of H1 require matching next-year audit opinions, which yields a prediction sample of 13,503 observations (fiscal years 2021-2023).

3.2 Tone Measurement

LLM tone (Tone_LLM). We use the DeepSeek LLM (deepseek-v4-flash). For each MD&A (first 30,000 characters, which concentrate the qualitative discussion), a uniform prompt instructs the model, acting as a senior financial analyst, to assess the overall tone of the managerial narrative (attending to how performance is characterized, how difficulties and risks are worded, and the confidence expressed about prospects) while explicitly disregarding the magnitude of reported financial figures. The model returns a continuous score from -1 (extremely pessimistic) to 1 (extremely optimistic). Temperature is set to zero for reproducibility.

Dictionary tone (Tone_Dict). We use the formal-text (annual report) sub-lexicon of the Chinese financial sentiment dictionary of Yao et al. [4] (3,592 positive and 1,633 negative words). After segmentation with jieba, we compute Tone_Dict = (Pos - Neg)/(Pos + Neg) for each MD&A.

3.3 Empirical Model

The baseline specification for H1 is:

$$ModOpinion(i,t+1) = a + b*Tone(i,t) + c*ModOpinion(i,t) + d'Controls(i,t) + Year\ \&\ Industry\ FE + e(i,t)$$

ModOpinion equals 1 for any opinion other than a standard unqualified opinion (unqualified with emphasis-of-matter paragraphs, qualified, or disclaimer) and 0 otherwise. Controls include firm size (log of total assets), leverage, ROA, revenue growth, inventory intensity, receivables intensity, operating cash flow scaled by assets, a loss indicator, Big 4 auditor, state ownership (SOE equals one if the ultimate controller is a central or local government agency, a state-owned assets supervision commission, or a state-owned enterprise, and zero otherwise; a small number of collectively-owned and other-type firms are classified as non-SOE), and ST status. Controlling for the current-year opinion isolates tone's incremental predictive content. We estimate logit models in the main tests and linear probability models (LPM) with industry and year fixed effects in robustness checks. Standard errors are clustered at the firm level throughout.

4. Empirical Results

4.1 Descriptive Statistics

Table 1 presents descriptive statistics. Tone_LLM has a mean of 0.449 and a median of 0.600, indicating that annual report narratives are systematically optimistic, consistent with prior evidence; its standard deviation of 0.333 reflects ample cross-sectional variation. Tone_Dict has a mean of 0.794 and a much tighter distribution (standard deviation 0.075), a first indication of the dictionary measure's limited discriminatory power. Modified opinions occur in 3.3% of firm-years; 20.9% of observations report losses; 28.0% are state-owned; 6.3% are audited by Big 4 firms. The correlation between the two tone measures is 0.396, suggesting they capture related but distinct textual attributes.

Table 1: Descriptive statistics

	N	Mean	SD	Min	P25	Median	P75	Max
Tone_LLM	18479	0.449	0.333	-0.95	0.2	0.6	0.7	1
Tone_Dict	18479	0.794	0.075	0.56	0.751	0.804	0.849	0.93
ModOpinion	18479	0.033	0.178	0	0	0	0	1
Ln(AuditFee)	18479	13.955	0.618	11.775	13.528	13.864	14.286	18.11
AuditLag (days)	18479	106.938	13.785	18	99	113	117	148
Size	18479	22.316	1.3	17.879	21.397	22.07	23.013	28.791
Leverage	18479	0.41	0.208	0.053	0.242	0.398	0.557	0.933
ROA	18479	0.03	0.071	-0.228	0.006	0.033	0.067	0.22
Growth	18479	0.098	0.31	-0.565	-0.071	0.063	0.214	1.527
Inventory	18479	0.122	0.102	0	0.052	0.101	0.161	0.556
Receivables	18479	0.129	0.101	0.001	0.05	0.108	0.186	0.454
OCF	18479	0.045	0.067	-0.157	0.008	0.045	0.084	0.24
Loss	18479	0.209	0.406	0	0	0	0	1

© 2026 The Author(s). Authors retain copyright in their articles.

	N	Mean	SD	Min	P25	Median	P75	Max
Big4	18479	0.063	0.242	0	0	0	0	1
SOE	18479	0.28	0.449	0	0	0	1	1
ST	18479	0.036	0.185	0	0	0	0	1

Notes: The sample comprises 18,479 firm-year observations of Chinese A-share listed firms over 2021-2024, excluding financial firms. Continuous variables are winsorized at the 1st and 99th percentiles. Variable definitions are provided in Section 3.

4.2 Tone Predicts Modified Audit Opinions (H1)

Table 2 reports the baseline results. In column (1), the logit coefficient on Tone_LLM is -0.618 (z = -3.15), significant at the 1% level: firms with more negative MD&A tone are more likely to receive a modified opinion the following year. Column (2) adds year and industry fixed effects, and column (3) switches to the LPM; the conclusion is unchanged (coefficients -0.705 and -0.019, both significant at 1%). Column (4) restricts the sample to firms with a clean current-year opinion, a pure "clean-to-modified" prediction test, and the coefficient strengthens to -0.820 (z = -3.47), indicating the result is not driven by opinion persistence among already-troubled firms.

In economic terms, the average marginal effect implied by column (1) is approximately -0.012: a one-unit increase in tone (about three standard deviations) lowers the probability of a next-year modified opinion by 1.2 percentage points, sizable against the 3.3% unconditional rate. Control variables behave as expected: leverage and ST status load positively; ROA, inventory intensity, and operating cash flow load negatively. H1 is supported.

Table 2: MD&A tone and next-year modified audit opinions

	(1) Logit	(2) Logit + FE	(3) LPM + FE	(4) Clean current opinion
Tone_LLM	-0.6184***	-0.7051***	-0.0187***	-0.8200***
Tone_LLM (t/z)	(-3.15)	(-3.54)	(-3.09)	(-3.47)
ModOpinion_t	3.2663***	3.2170***	0.4367***	
ModOpinion_t (t/z)	(17.69)	(17.24)	(18.09)	
Size	0.0383	0.0550	0.0004	0.0601
Size (t/z)	(0.62)	(0.90)	(0.31)	(0.84)
Leverage	0.9816***	1.1825***	0.0346***	2.0105***
Leverage (t/z)	(2.83)	(3.22)	(3.17)	(4.57)
ROA	-2.6971**	-2.3023*	-0.0869**	-4.2988***
ROA (t/z)	(-2.14)	(-1.79)	(-2.06)	(-2.65)
Growth	0.4365**	0.4363**	0.0139**	0.2559
Growth (t/z)	(2.23)	(2.12)	(2.29)	(1.09)
Inventory	-1.6466***	-2.2038***	-0.0621***	-2.1879***
Inventory (t/z)	(-2.69)	(-3.38)	(-4.34)	(-2.88)
Receivables	0.4328	0.5997	0.0123	0.6447
Receivables (t/z)	(0.74)	(0.92)	(0.77)	(0.86)
OCF	-2.3291**	-2.5254**	-0.0298	-3.1152**
OCF (t/z)	(-2.38)	(-2.47)	(-1.16)	(-2.49)
Loss	0.3277	0.3060	0.0063	0.1356
Loss (t/z)	(1.52)	(1.41)	(1.14)	(0.52)
Big4	-0.4325	-0.4957	-0.0055	-0.2858
Big4 (t/z)	(-1.23)	(-1.35)	(-1.33)	(-0.77)
SOE	-0.3472**	-0.3409**	-0.0089***	-0.6477***
SOE (t/z)	(-2.12)	(-2.00)	(-2.78)	(-3.29)
ST	2.3522***	2.3416***	0.2243***	2.6280***
ST (t/z)	(14.87)	(14.74)	(11.07)	(13.84)
Year & Industry FE	No	Yes	Yes	No
N	13503	13503	13503	13066
Adj. R2 / Pseudo R2	0.417	0.426	0.342	0.214

*Notes: The dependent variable is ModOpinion(t+1). Columns (1), (2), and (4) report logit estimates; column (3) reports the linear probability model. Column (4) restricts the sample to firms with a standard unqualified opinion in year t. t/z statistics based on firm-clustered standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

© 2026 The Author(s). Authors retain copyright in their articles.

4.3 Mechanism: Forward-Looking Information in Tone (H2)

Why does tone predict opinions? Our explanation is that tone reflects managers' private information about the firm's true condition; that information subsequently materializes in fundamentals, to which the auditor's risk assessment responds. Table 3 provides direct evidence. Controlling for current ROA, Tone_LLM predicts next-year ROA with a coefficient of 0.013 (t = 7.37) and next-year loss incidence with a logit coefficient of -0.668 (z = -8.02), both significant at the 1% level. Optimistic-tone firms genuinely perform better in the following year; negative-tone firms are more likely to slide into losses. H2 is supported.

Table 3: Forward-looking information content of tone

	(1) ROA t+1 (OLS)	(2) Loss t+1 (Logit)
Tone_LLM	0.0134***	-0.6682***
Tone_LLM (t/z)	(7.37)	(-8.02)
ROA	0.5645***	-12.1438***
ROA (t/z)	(33.91)	(-12.68)
Loss	-0.0031	0.9970***
Loss (t/z)	(-1.57)	(10.19)
Year & Industry FE	Yes	No
N	13503	13503
Adj. R2 / Pseudo R2	0.463	0.252

*Notes: Column (1) reports OLS estimates with ROA(t+1) as the dependent variable; column (2) reports logit estimates with Loss(t+1) as the dependent variable. t/z statistics based on firm-clustered standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

4.4 Horse Race: LLM Tone Versus Dictionary Tone (H3)

Table 4 pits the two measures against each other. Used alone, Tone_Dict does not predict next-year modified opinions (coefficient -0.159, z = -0.19); used alone, Tone_LLM does. With both included, Tone_LLM remains significant at the 1% level (-0.707, z = -3.23) while Tone_Dict remains insignificant. The audit-relevant component of tone is therefore precisely the contextual, semantic component that word frequencies cannot capture; it is recoverable only by a reader, human or machine, that understands language in context. H3 is supported, and the finding shows that our results could not have been obtained with the conventional method.

Table 4: Horse race between LLM tone and dictionary tone

	(1) LLM only	(2) Dictionary only	(3) Both
Tone_LLM	-0.6184***		-0.7073***
Tone_LLM (t/z)	(-3.15)		(-3.23)
Tone_Dict		-0.1590	1.0119
Tone_Dict (t/z)		(-0.19)	(1.04)
Year & Industry FE	No	No	No
N	13503	13503	13503
Adj. R2 / Pseudo R2	0.417	0.414	0.417

*Notes: The dependent variable is ModOpinion(t+1); all columns report logit estimates with the full control set of Table 2. t/z statistics based on firm-clustered standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

4.5 Validity of the LLM Measure

4.5.1 Look-Ahead Bias: An Anonymization Test

The central identification concern for LLM-based measures is look-ahead bias: the model may have "seen" these firms' subsequent outcomes in its training corpus, so its scores could encode memory rather than reading [21,22]. To address this, we draw 2,000 firm-year observations (all 446 next-year-modified cases plus a random control group), strip each MD&A of all identifying information (company names, name variants, and security codes are replaced with generic placeholders), and have the model rescore the anonymized texts under the identical prompt. Table 5 shows that anonymized tone correlates 0.788 (Pearson; 0.787 Spearman) with the original scores, with nearly identical means and standard deviations. More importantly, within this subsample the anonymized tone retains significant predictive power for next-year modified opinions (coefficient -0.668, z = -2.44), close in direction and magnitude to the original measure (-0.855, z = -3.32).

© 2026 The Author(s). Authors retain copyright in their articles.

The model identifies audit risk from the text alone, without knowing which firm it is reading. Look-ahead bias cannot explain our findings. The slightly larger coefficient on the original score admits the possibility that firm identity contributes a modest increment of information, but the predictive power of tone derives predominantly from the text itself.

Table 5: Anonymization test for look-ahead bias

Item	Value	
Matched observations	1996	
Pearson corr. (original vs. anonymized)	0.788	
Spearman corr.	0.787	
Original tone: mean / SD	0.398 / 0.370	
Anonymized tone: mean / SD	0.426 / 0.368	
	(1) Original tone	(2) Anonymized tone
Tone_LLM	-0.8548***	
Tone_LLM (t/z)	(-3.32)	
Tone_Anon		-0.6676**
Tone_Anon (t/z)		(-2.44)
Year & Industry FE	No	No
N	1996	1996
Adj. R2 / Pseudo R2	0.476	0.474

*Notes: The anonymization subsample comprises 1,996 firm-year observations (all next-year-modified cases plus a random control group). The lower panel reports logit estimates with ModOpinion(t+1) as the dependent variable and the full control set of Table 2. t/z statistics based on firm-clustered standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

4.5.2 Human Validation

We draw a stratified random sample of 100 MD&A texts (20 from each quintile of the LLM tone distribution, order randomized, all firm identifiers withheld) and have two trained coders independently and blindly classify each text as optimistic, neutral, or pessimistic under a common coding manual. We follow a two-stage protocol. In the first stage, independent coding yields 51% raw agreement (Cohen's kappa = 0.02), with disagreements concentrated on the neutral-optimistic boundary, an indication of how difficult tone judgment is for non-expert readers, and itself a motivation for machine-based measurement. In the second stage, the coding manual is refined to resolve the sources of disagreement (separating firm-authored narrative from quoted industry statistics; classifying "difficulty-response-achievement" framing), and the two coders jointly adjudicate each disputed case to a consensus label, with a handful of deadlocked cases resolved by a third adjudicator applying the manual. The final consensus labels comprise 57 optimistic and 43 neutral texts (no pessimistic labels, consistent with the well-documented positive skew of annual report narratives). Validation results are strong: mean LLM tone is 0.594 for texts labeled optimistic versus 0.355 for neutral ($t = 4.37$, $p < 0.001$), and the Spearman rank correlation between LLM scores and consensus labels is 0.442 ($p < 0.001$). The LLM's scores align significantly with vetted human judgment, providing direct evidence of measurement validity.

4.6 A Contrasting Result: Tone, Audit Fees, and Audit Report Lag

Prior U.S. studies typically use audit fees as the outcome [15]. Table 6 shows that in our sample the univariate relation between tone and log audit fees is significantly positive, but the coefficient collapses to zero once firm size and other controls are added (-0.003, $t = -0.19$). Tone is likewise unrelated to audit report lag ($t = -0.75$). Alternative specifications, including abnormal tone (the residual from regressing tone on fundamentals), first differences, firm fixed effects, and lagged tone, yield the same null (available on request). This null result is informative when juxtaposed with the opinion results: audit fees in China are sticky and dominated by hard client characteristics such as size, leaving little room for soft information in pricing; yet auditors' risk judgments, expressed in opinions, do impound tone. Soft information shapes what auditors conclude, not what they charge.

Table 6: Tone, audit fees, and audit report lag

	(1) Ln(AuditFee), FE only	(2) Ln(AuditFee), full controls	(3) AuditLag, full controls
Tone_LLM	0.1021***	-0.0025	-0.2959

© 2026 The Author(s). Authors retain copyright in their articles.

Tone_LLM (t/z)	(5.55)	(-0.19)	(-0.75)
Year & Industry FE	Yes	Yes	Yes
N	18479	18479	18479
Adj. R2 / Pseudo R2	0.099	0.605	0.105

*Notes: The dependent variable is the natural logarithm of audit fees in columns (1)-(2) and audit report lag (days from fiscal year-end to the audit report date) in column (3). All columns include year and industry fixed effects. t statistics based on firm-clustered standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

4.7 Which Type of Modified Opinion Does Tone Predict?

A modified opinion in China spans two economically distinct categories: an unqualified opinion with an emphasis-of-matter paragraph, and a severe modification (qualified opinion, disclaimer, or adverse opinion). The two arise for different reasons. An emphasis-of-matter paragraph is the auditor's device for flagging going-concern doubt or significant uncertainty without qualifying the opinion, precisely the kind of soft, forward-looking risk that managerial tone would reveal. Severe modifications, by contrast, stem from concrete GAAP departures or scope limitations, that is, hard reporting failures with no intrinsic connection to how management narrates the year. We therefore expect tone to predict the former but not the latter.

Table 7 confirms this. Using next-year emphasis-of-matter (1.9% of firm-years) as the outcome, Tone_LLM is strongly predictive (coefficient -0.789, z = -3.45, p < 0.001); using next-year severe modifications (1.4% of firm-years), tone has no predictive power (coefficient -0.142, z = -0.49). The baseline result is thus driven entirely by emphasis-of-matter paragraphs. This decomposition sharpens the interpretation: tone predicts auditors' concern about uncertainty, not their detection of hard reporting failures, consistent with our thesis that tone conveys soft, forward-looking information about the firm's condition.

Table 7: Which type of modified opinion does tone predict?

	(1) Emphasis-of-matter	(2) Severe modification
Tone_LLM	-0.7890***	-0.1419
Tone_LLM (t/z)	(-3.45)	(-0.49)
ModOpinion_t	2.7008***	2.4406***
ModOpinion_t (t/z)	(10.31)	(9.86)
Year & Industry FE	No	No
N	13503	13503
Adj. R2 / Pseudo R2	0.329	0.326

*Notes: The dependent variable is an indicator for a next-year unqualified opinion with an emphasis-of-matter paragraph in column (1), and an indicator for a next-year severe modification (qualified, disclaimer, or adverse) in column (2). Both columns report logit estimates with the full control set of Table 2. t/z statistics based on firm-clustered standard errors are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

5. Robustness Tests

In addition to the checks reported above, our results are robust to : (i) excluding 2021 or any single year from the sample; (ii) replacing the continuous tone measure with an indicator for tone below the industry-year median; and (iii) augmenting the mechanism regressions with additional controls for size and leverage.

6. Conclusion

Using an LLM to measure the tone of roughly 20,000 MD&A sections of Chinese listed firms over 2021-2024, we show that narrative tone negatively predicts next-year modified audit opinions; that this effect is driven by emphasis-of-matter paragraphs, the auditor's signal of going-concern doubt and uncertainty, rather than by severe modifications rooted in hard reporting failures; that it operates through forward-looking information about fundamentals; that the predictive component of tone is invisible to a state-of-the-art sentiment dictionary; and that an anonymization experiment rules out look-ahead bias. Tone does not, in contrast, explain audit fees or report lag.

The implications are threefold. For researchers, LLMs provide a measurement technology for (Chinese) financial text that materially outperforms dictionaries, but LLM-based measures should be accompanied by

© 2026 The Author(s). Authors retain copyright in their articles.

validity checks, and we offer a replicable two-part framework (dictionary horse race plus anonymization test). For auditors and regulators, narrative tone is an inexpensive, machine-readable leading indicator of audit risk; firms whose tone is abnormally negative, or inconsistent with reported fundamentals, warrant incremental scrutiny, and exchanges could incorporate tone screening into inquiry-letter targeting. For investors, how management speaks contains information beyond what it reports, and is worth integrating into fundamental analysis.

Limitations suggest avenues for future work. Our sample period is short, constrained by text availability; scores depend on a specific model and prompt, and stability across models and prompts merits systematic study; and the method extends naturally to earnings-call transcripts, inquiry-letter responses, and other corporate texts.

References

[1] Li F. Annual report readability, current earnings, and earnings persistence[J]. Journal of Accounting and Economics, 2008, 45(2-3):221-247.

[2] Loughran T, McDonald B. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks[J]. Journal of Finance, 2011, 66(1):35-65.

[3] Davis A K, Piger J M, Sedor L M. Beyond the numbers: Measuring the information content of earnings press release language[J]. Contemporary Accounting Research, 2012, 29(3):845-868.

[4] Yao J, Feng X, Wang Z, et al. Tone, sentiment and market impacts: The construction of a Chinese sentiment dictionary in finance[J]. Journal of Management Sciences in China, 2021, 24(5):26-46. (In Chinese)

[5] Li F. The information content of forward-looking statements in corporate filings—A naive Bayesian machine learning approach[J]. Journal of Accounting Research, 2010, 48(5):1049-1102.

[6] Huang X, Teoh S H, Zhang Y. Tone management[J]. The Accounting Review, 2014, 89(3):1083-1113.

[7] DeFond M L, Zhang J. A review of archival auditing research[J]. Journal of Accounting and Economics, 2014, 58(2-3):275-326.

[8] Loughran T, McDonald B. Textual analysis in accounting and finance: A survey[J]. Journal of Accounting Research, 2016, 54(4):1187-1230.

[9] Kim A, Muhn M, Nikolaev V. Bloated disclosures: Can ChatGPT help investors process information?[R]. Working paper, University of Chicago, 2023.

[10] Lopez-Lira A, Tang Y. Can ChatGPT forecast stock price movements? Return predictability and large language models[R]. Working paper, University of Florida, 2023.

[11] Feldman R, Govindaraj S, Livnat J, et al. Management's tone change, post earnings announcement drift and accruals[J]. Review of Accounting Studies, 2010, 15(4):915-953.

[12] Xie D, Lin L. Can management tone predict firms' future performance? Evidence from the earnings communication conferences of Chinese listed companies[J]. Accounting Research, 2015(2):20-27. (In Chinese)

[13] Zeng Q, Zhou B, Zhang C, et al. Annual report tone and insider trading: Consistent words and deeds or saying one thing and doing another?[J]. Management World, 2018, 34(9):143-160. (In Chinese)

[14] Simunic D A. The pricing of audit services: Theory and evidence[J]. Journal of Accounting Research, 1980, 18(1):161-190.

[15] Abernathy J L, Guo F, Kubick T R, et al. Financial statement footnote readability and corporate audit outcomes[J]. Auditing: A Journal of Practice & Theory, 2019, 38(2):1-26.

[16] DeFond M L, Wong T J, Li S. The impact of improved auditor independence on audit market concentration in China[J]. Journal of Accounting and Economics, 2000, 28(3):269-305.

© 2026 The Author(s). Authors retain copyright in their articles.

[17] Chen C J P, Su X, Zhao R. An emerging market's reaction to initial modified audit opinions: Evidence from the Shanghai Stock Exchange[J]. Contemporary Accounting Research, 2000, 17(3):429-455.

[18] Chen C J P, Chen S, Su X. Profitability regulation, earnings management, and modified audit opinions: Evidence from China[J]. Auditing: A Journal of Practice & Theory, 2001, 20(2):9-30.

[19] Hu N, Xu J, Xue S. Regulatory risk and auditors' reporting conservatism: Evidence from Chinese comment letters[J]. Journal of Accounting and Public Policy, 2022, 41:106997.

[20] Kim A, Muhn M, Nikolaev V. Financial statement analysis with large language models[R]. Working paper, University of Chicago, 2024.

[21] Glasserman P, Lin C. Assessing look-ahead bias in stock return predictions generated by GPT sentiment analysis[R]. Working paper, Columbia University, 2023.

[22] Sarkar S K, Vafa K. Lookahead bias in pretrained language models[R]. Working paper, Harvard University, 2024.

Funding

This research was supported by the College Students' Innovation and Entrepreneurship Training Program of Nanjing Audit University [grant number: X2025112870050].

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.8496
Published by: Zeus Press

Cross-Sectional Return Prediction in China's A-Share Market Based on Lasso and XGBoost: A Comparison with the Fama–MacBeth Multi-Factor Model

Xichun Zhang*

School of Mathematics and Statistics, North China University of Water Resources and Electric Power, Zhengzhou 450046, China

**Corresponding author: Xichun Zhang.*

Abstract

In response to the problem of the widespread use of mixed prediction performance metrics in machine learning stock selection literature, the lack of empirical tests on the applicability boundaries of linear and tree-based models under low-dimensional factor settings, and the disconnection of most high-dimensional factor empirical conclusions from the actual scenarios of small and medium-sized quantitative institutions, this paper selects the monthly panel data of A-shares from 2013 to 2023 to avoid the leakage of look-ahead information, and builds a unified rolling-window forecasting framework of FM, Lasso-FM, and XGBoost. The model hyperparameters are selected based on the existing literature to ensure a fair comparison among the three models. It relies on OOS-R^2, IC, Rank-IC and long-short portfolio Sharpe ratio to carry out multi-dimensional evaluation, and carries out analysis in combination with the heterogeneity of market value grouping and SHAP decomposition. In addition, multiple sets of robustness checks are conducted. The empirical evidence shows that under the sample setting of the five basic factors in this article, the mean out-of-sample OOS- R^2 values of all three models are negative, the absolute return prediction performance is weak, and there is no significant statistical superiority or inferiority between the models; However, there is a deviation in the evaluation index, and XGBoost performs better in cross-sectional stock ranking and long–short portfolio performance, which is rooted in the fact that OOS-R^2 focuses on absolute error, and IC only focuses on the relative return strength of stocks; Based on the observation of this sample, it is difficult for XGBoost to release nonlinear modeling capability, and greater model flexibility has the possibility of inducing out-of-sample overfitting. The forecast stability of large-cap stocks is significantly higher than that of small-cap stocks, and the market value and ROA are important prediction features of the model. This paper explains the intrinsic mechanism of multi-indicator deviation, defines the applicable boundaries of factor dimensionality of different models, and provides a reference for target lightweight modeling for small and medium-sized institutions.

Keywords

cross-sectional return prediction, Fama–MacBeth, Lasso, XGBoost, SHAP, out-of-sample OOS- R^2

1. Introduction

Machine learning methods are widely used in the cross-sectional return prediction research of A-share stocks, providing new analytical tools for quantitative stock selection. Extreme Gradient Boosting (XGBoost)

© 2026 The Author(s). Authors retain copyright in their articles.

shows a good predictive effect in a large number of high-dimensional empirical studies with the ability to mine the high-order interaction of factors [1,2,3]. Most of the existing domestic related studies are based on the hundred-dimensional derived factors to carry out backtesting to verify the out-of-sample prediction advantages of XGBoost. However, the prerequisites of this kind of research deviate from the actual modeling scenario of small and medium-sized institutions in China. Small and medium-sized institutions can often only obtain a small number of basic factors such as financial and market variables, and it is difficult for the research conclusions obtained in the high-dimensional environment to be directly transferred to low-dimensional modeling practice [4]. At the same time, some literature often mixes many indicators such as out-of-sample R^2, OOS-R^2, Information Coefficient (IC), and rank IC in the evaluation of the model performance. The indicator divergence phenomenon of ”out-of-sample absolute fitting fails, but the stock portfolio can still obtain positive returns“ has not been systematically explained; Most studies are based on high-dimensional data sets to derive the judgment that the nonlinear model is comprehensively superior, but rarely test whether the conclusion is valid under the condition of low-dimensional basic factors, and lack the normative empirical test of the applicability boundaries of linear and tree-based models [5]. In addition, some similar studies also have empirical problems such as look-ahead information leakage, inconsistence of multi-model comparison benchmarks, and confusion between the internal prediction weight of the model and the factor risk premium.

In view of the above research shortcomings, this article aims to clarify the inherent logical differences of different prediction and evaluation indicators, explore the relative performance of linear and machine learning models under the constraints of low-dimensional basic factors, and supplement the empirical experience of A-share low-dimensional scenarios. This article selects the monthly panel data of A-shares from 2013-2023, uses the design of t-1 period public factor delay to predict returns at time t+1 to avoid look-ahead bias, builds three types of prediction models of Fama-MacBeth (Fama-MacBeth, FM), Least Absolute Shrinkage and Selection Operator (Lasso-FM), and XGBoost, and builds a unified rolling-window backtesting framework. The model hyperparameters are selected based on the existing literature to select fixed values to ensure the fairness of the comparison of the three types of models. In combination with SHapley Additive exPlanations (SHAP), we carry out factor contribution analysis, and at the same time, we combine the circulation market value grouping for heterogeneity analysis, and set up multiple groups of independent robust tests to improve the reliability of the results, in order to provide reference ideas for lightweight modeling for small and medium-sized institutions with limited computing power and factor resources. It should be noted that the empirical conclusion of this article only adapts to the sample and factor settings of this study, and is not universal.

2. Literature Review and Theoretical Framework

2.1 Literature Review

The classical linear analysis framework of asset pricing includes CAPM, Fama-French factor model, and Fama-MacBeth cross-section regression. Previous international studies have shown that under the massive supply of features, XGBoost relies on high-order nonlinear interaction to improve the ability to predict out-of-sample returns, but the gain is highly dependent on sufficient factor input [6]. There is an obvious division in the relevant domestic research. Wu Shinong and Xu Nianxing believe that the speculative behavior of A-share retail investors raises the monthly return noise, and the high-degree-of-freedom nonlinear model is easy to fit random fluctuations with no economic meaning [1]. Li Bin verified the data noise reduction effect of L1 regularization, but did not make a parallel comparison with XGBoost [4]. Wang Xi and others only test the low-dimensional Lasso model, and lack the machine learning reference benchmark [7]. Gu Songnan, Qiao Zheng and others rely on the high-dimensional derivative factor to draw the superior conclusion of XGBoost, which is difficult to adapt to the low-factor practical scenarios of small and medium-sized institutions [2,3]. A domestic review has pointed out that machine learning can improve multi-factor pricing in a big data environment, but the comparative empirical evidence of linear and tree models under low-dimensional basic factor conditions is still insufficient [5]; The existing autoencoder-based machine learning asset pricing research points out that the strong fitting ability of machine learning is easier to capture market noise, and the model’s predictive performance is limited under low-dimensional conditions [8]. Herd and noise trading are important features of the A-share market, which will amplify the risk of model overfitting [9]; the empirical evidence of deep learning and tree models in the domestic market is mostly based on high-dimensional feature data sets [10,11]. The existing literature lacks parallel comparison of low-dimensional basic factors under a

© 2026 The Author(s). Authors retain copyright in their articles.

unified framework, and the interpretation of the economic transmission mechanism of empirical results is insufficient, which is supplemented by this article.

2.2 Core Theoretical Mechanisms

2.2.1 Fundamental Logic of Factor-Based Prediction

The five mature forecast indicators of log market capitalization LnSize, book market value ratio BM, return on assets (ROA), monthly turnover rate and monthly momentum are selected, which correspond to the five return-generating mechanisms of liquidity premium, valuation repair, sustainable profit, short-term speculation and price trend respectively. Extensive evidence from China has demonstrated the effectiveness of these factors in monthly return prediction.

2.2.2 Applicability Boundaries of Linear and Tree-Based Models

FM and Lasso-FM only capture factor linear association; The core value of XGBoost is to mine multi-factor high-order interactions, and this ability is only fully released in the hundred-dimensional high-factor pool. Under the sample setting of the five basic factors in this article, there are very few interactive relationships that can be mined, and the inherent advantages of the tree model are difficult to give full play. At the same time, the monthly return noise of A-shares is high, the XGBoost model flexibility is higher, and is more likely to memorize random noise in the training data, leading to overfitting and weaker out-of-sample predictive performance.

Traditional Lasso is mostly used to screen redundant variables in high-dimensional data sets. There are no large number of redundant features under low-dimensional conditions in this article. The L1 regularization no longer assumes the screening function. It only slightly alleviates the mild multicollinearity of LnSize and BM, compresses the low contribution coefficient of noise-driven, and improves the stability of model generalization. The full text weakens the relevant strong expression of “Lasso high-dimensional screening, XGBoost complex nonlinear mining”, which fits the five-factor sample constraints.

2.2.3 Potential Channels of Heterogeneous Predictive Performance across Market Capitalization Groups

Large-market individual stock institutions hold positions, analyst tracking coverage is sufficient, information disclosure is standardized, and fundamental forecast signals are less interfered by short-term speculation; Small-sized markets are mainly retail transactions, and the subject speculation is frequent. Superimposed information disclosure is lagging behind, institutional coverage is insufficient, forming information asymmetry, and effective income signals are easily diluted by random capital fluctuations. This article does not introduce institutional shareholding and analyst coverage agent variables to carry out intermediary tests. The stratification of large and small plate indicators is only for correlation observation, and information asymmetry is only a potential explanation path, without causal judgment. The relevant intermediary conduction is left for further expansion.

2.2.4 Mechanism of Rolling Window Constraints

The benchmark rolling window is set to 60 months, taking into account the sample capacity and data timeliness; the 36- and 48-month windows are only used for model hyperparameter sensitivity testing, and are not standardized independent robust testing.

3. Research Design

3.1 Data and Variable Definitions

The sample period covers monthly panel data for A-shares from CSMAR, spanning January 2013 to December 2023. Data preprocessing rules: Exclude ST, *ST, and delisting-watch stocks; remove samples with missing data for the factor at period t or the return at period t+1. Individual stock excess returns are calculated using the 3-month time deposit interest rate; extreme values for variables such as return, turnover rate, and momentum are truncated, and all continuous variables undergo two-sided winsorization (at the 1% and 99% levels) to eliminate extreme outliers.

© 2026 The Author(s). Authors retain copyright in their articles.

a) Dependent variable: $Mret_{i,t+1}$: The raw monthly return of stock *i* for month *t*+1. Throughout the process, factors from period *t* are used with a one-period lag to predict returns for period *t*+1; using contemporaneous factors to explain contemporaneous returns is prohibited to avoid look-ahead bias and information leakage.

b) Predictors (all based on month-end observations for period *t*): Log of market capitalization (LnSize), book-to-market ratio (BM), return on assets (ROA), monthly turnover rate, and monthly momentum.

c) Grouping variable: Month-end tradable market capitalization of individual stocks; stocks are classified into large-cap and small-cap groups based on the monthly sample median.

3.2 Model Specification

The model inputs, prediction targets, and time-series partitioning are completely consistent across all three types of models, ensuring a fair comparison.

3.2.1 Traditional Fama–MacBeth Cross-Sectional Regression Model

Cross-sectional linear regressions of individual stocks are estimated monthly, and the resulting time-varying coefficients are used to generate return forecasts for the following month:

$$Mret_{i,t+1} = \alpha_t + \sum_{j=1}^{J} \beta_{j,t}\, Factor_{i,j,t} + \varepsilon_{i,t+1} \tag{1}$$

Where: $Mret_{i,t+1}$ is the raw monthly return of stock *i* for month *t*+1; α_t is the monthly intercept term; $\beta_{j,t}$ is the cross-sectional regression coefficient for the *j*-th factor in period *t*; $Factor_{i,j,t}$ is the value of the *j*-th factor for stock *i* in period *t*; $\varepsilon_{i,t+1}$ is the random disturbance term; and $\sum$ is the summation operator.

3.2.2 Lasso–Fama–MacBeth Linear Shrinkage Model

Introducing L1 regularization into the linear least-squares loss function shrinks the coefficients of low-contribution factors and mitigates multicollinearity:

$$\min_{\beta} \sum_{i=1}^{N} \quad (Mret_{i,t+1} - \alpha_t - \sum_{j=1}^{J} \beta_{j,t}\, Factor_{i,j,t})^2 + \lambda \sum_{j=1}^{J} |\beta_{j,t}|\# \tag{2}$$

In the equation, *min* denotes the minimization operation; λ is the L1 regularization penalty coefficient; and the other symbols have the same meanings as in Equation (1). This paper selects a fixed penalty coefficient for modeling, drawing upon existing quantitative research on A-shares.

3.2.3 XGBoost Nonlinear Ensemble Prediction Model

Using five factors from period t as inputs, multiple decision trees are iteratively fitted to model the return for period *t*+1.

$$y_{i,t+1} = \sum_{k=1}^{K} f_k\,(Factor_{i,t}) \tag{3}$$

Where $y_{i,t+1}$ is the predicted return of stock *i* for month *t*+1; $f_k(\cdot)$ is the *k*-th decision tree model; *K* is the total number of decision trees; $\sum$ is the summation symbol. This paper sets the model's fixed hyperparameters based on similar literature; under low-dimensional constraints, only simple factor interactions can be captured.

3.3 Standardized Rolling-Window Procedure

This paper employs a rolling window approach for out-of-sample forecasting. The benchmark regression utilizes a fixed 60-month training window that rolls forward month by month. For each period, the model is trained using the full history of monthly samples within the rolling window to forecast individual stock returns for the subsequent month; this strictly adheres to the principle of using factors from period t to predict returns for period t+1, thereby avoiding look-ahead bias. The hyperparameters for Lasso and XGBoost were selected based on fixed values from domestic A-share return prediction literature; no independent validation set was used for grid search. To examine the sensitivity of the conclusions to the window length, the robustness check

© 2026 The Author(s). Authors retain copyright in their articles.

includes two additional sets of rolling windows—48 months and 36 months—using which the aforementioned backtesting process is repeated.

Fixed model parameters: Lasso penalty coefficient $\alpha = 0.0001$; XGBoost number of trees n_estimators = 100, maximum tree depth max_depth = 3. Ordinary Least Squares (OLS) estimation is used for the Fama-MacBeth procedure.

3.4 Multidimensional Evaluation Framework for Predictive Performance

This article adopts four-layer evaluation standards and discards the single OOS-R^2 indicator: (1) Out-of-sample fitting: monthly OOS-R^2; (2) Sorting index: information coefficient IC, Rank Information Coefficient Rank IC; (3) Portfolio performance: decile long–short portfolio returns, annualized Sharpe ratio, deduct 0.3% bilateral cost income; (4) Statistical test: A Newey–West adjusted *t*-test is conducted on the monthly differences in the performance metrics to determine whether the differences between models are statistically significant.

3.5 Distinction between the Validation Frameworks

Distinguish between window sensitivity and independent robustness test to avoid conceptual confusion: window adjustment only verifies the tolerance of the model to parameters, and does not change the core settings of samples and factors; The five types of robust tests change the research settings from the multi-dimensionality of sample, return definition, factor construction, benchmark model and transaction cost, which is used to prove the universality of the conclusion, and the two tests are logically independent.

a) Window sensitivity test: 36, 48-month rolling length, only for parameter endurance verification, not a robust test.

b) Five types of independent robustness tests: segmented sub-sample, excess return substitution, factor industry neutrality, gradual regression benchmark, and multi-stage transaction cost.

Expansion analysis: the important prediction features of market value grouping heterogeneity and SHAP global average absolute value quantitative variables, strictly distinguish between model fitting weight and long-term factor pricing premium, and the two main lines of analysis should not be confused with each other.

4. Empirical Results and Analysis

4.1 Descriptive Statistics of the Main Variables

The effective observation of the whole sample is a total of 117,540 groups, with the average return of individual shares in the next month -0.0067, and the standard deviation is 0.1313, which visually reflects the high noise characteristics of the monthly return of A-shares; The deviation of the original extreme value of the monthly change rate and the monthly momentum is extremely large. This article standardizes preprocessing all core variables: first, the upper and lower limits of the variables with extreme change, such as monthly return, change of hands, momentum, etc., are cut off, and the extreme error values that do not conform to the realistic logic of the A-share market are eliminated; Subsequently, 1% and 99% percential bilateral shrinkage treatment is carried out on all variables to further weaken the interference of extreme observation on model estimation and ensure the robustness of empirical results.

Table 1: Descriptive statistics of main variables

Variable	Observations	Mean	Standard Deviation	Minimum	Median	Maximum
Next-month stock return	117540	-0.0066	0.1311	-0.2987	-0.0171	0.5843
Log market capitalization	117540	15.7776	1.1144	12.3869	15.6437	18.9772
Book-to-market ratio	117540	0.4154	0.2205	0.0799	0.4008	20.5246
Return on assets (ROA)	117540	0.0574	0.1338	-0.1320	0.0462	0.1408
Monthly turnover rate	117540	0.0015	0.0194	-0.0065	0.0003	1.0000
Monthly momentum	117540	0.0082	0.0636	-0.0244	-0.0006	1.0000

Note: This table reports the descriptive statistics of the variables after outlier treatment and two-sided winsorization at the 1st and 99th percentiles. All variables are reported in decimal form, and their distributions are broadly consistent with the typical characteristics of the China A-share market.

© 2026 The Author(s). Authors retain copyright in their articles.

4.2 Preliminary Test for Multicollinearity among Factors

This study employs the Pearson correlation coefficient and the Variance Inflation Factor (VIF) to assess multicollinearity among the variables; the maximum pairwise correlation coefficient between factors is only 0.1928, and the VIF values for all factors are below 10. Only the logarithm of market capitalization and the book-to-market ratio exhibit mild multicollinearity; no multicollinearity is present among the other variables, allowing them to be directly incorporated into the three types of models for regression fitting.

Table 2: Pearson Correlation Coefficient Matrix and VIF Collinearity Test

Factor	Book-to-Market Ratio	Return on Assets (ROA)	Monthly Turnover Rate	Log Market Capitalization	Monthly Momentum	VIF	Collinearity Assessment
Book-to-market ratio	1.0000	-0.1262	0.0361	0.1928	0.0068	4.8112	Mild multicollinearity
Return on assets (ROA)	-0.1262	1.0000	0.0444	0.0745	0.0260	1.2160	No significant multicollinearity
Monthly turnover rate	0.0361	0.0444	1.0000	0.0045	0.1723	1.0404	No significant multicollinearity
Log market capitalization	0.1928	0.0745	0.0045	1.0000	-0.0061	5.2254	Mild multicollinearity
Monthly momentum	0.0068	0.0260	0.1723	-0.0061	1.0000	1.0480	No significant multicollinearity

Notes: (a) The first five columns report the pairwise Pearson correlation coefficients among the factors. (b) Following the commonly adopted criterion, a VIF value below 10 indicates the absence of severe multicollinearity. (c) All statistics were computed in batches using the statsmodels package in Python.

4.3 Monthly Coefficient Estimates of the Linear Models

The LnSize coefficient in the FM model is significantly negative, consistent with established domestic and international findings regarding the size premium [12]; ROA exhibits the strongest linear predictive power. The unconstrained FM lacks a regularization term and is prone to fitting noise; while Lasso significantly shrinks factor coefficients, their signs remain unchanged. This indicates that regularization suppresses random disturbances without altering the direction of factor return predictions, thereby reducing noise and enhancing generalization capability.

Table 3: Average monthly regression coefficients

Factor	Average FM Coefficient	Average Lasso–FM Coefficient
Log market capitalization	-0.000265	-0.000108
Book-to-market ratio	0.006044	0.001740
Return on assets (ROA)	0.019875	0.000840
Monthly turnover rate	0.000719	0.000438
Monthly momentum	0.000055	0.000059

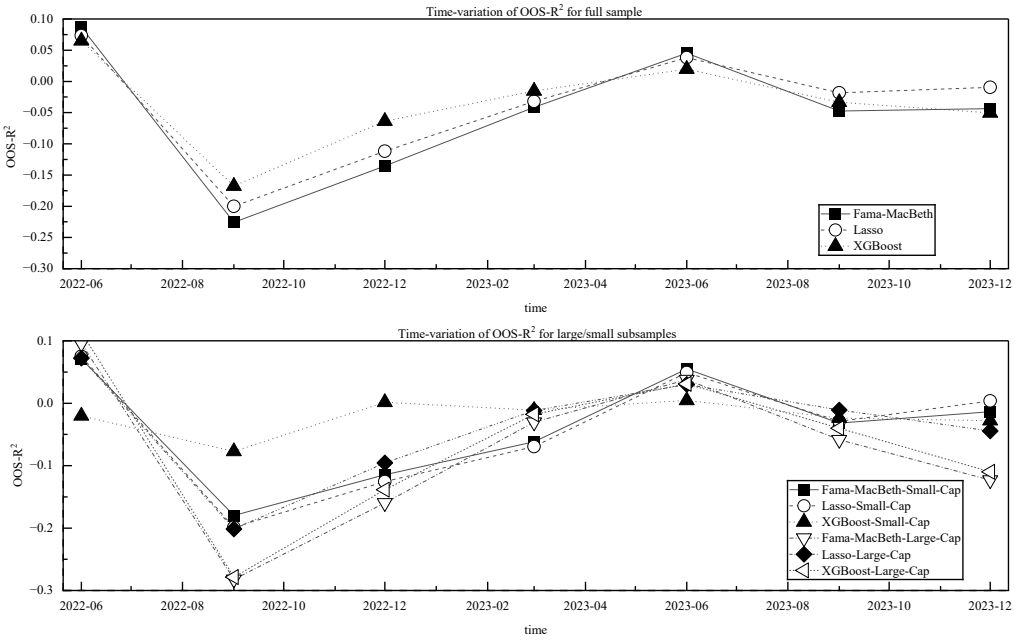
4.4 Multidimensional Prediction Results under the Benchmark Window and Analysis of Market Capitalization Heterogeneity

4.4.1 Overall Predictive Performance under Different Rolling Window Lengths

As shown in Figure 1, under the 60-month benchmark window, the average values of OOS-R^2 of FM, Lasso-FM and XGBoost are -0.0241, -0.0097 and -0.0328 respectively. The out-of-sample decision coefficients of the three types of models are all negative, indicating that the absolute level prediction ability of the monthly return of individual stocks by relying on the five basic factors of this group is limited. Although the OOS-R^2 mean of Lasso-FM is relatively higher, the Newey-West robust t-test shows that there is no statistically significant difference between the models, and it cannot be determined that Lasso-FM has a deterministic advantage. At the same time, the indicators are differentiated: XGBoost's IC, Rank IC, the annualized Sharpe ratio of the decile long–short portfolio and other sorting and Portfolio performance indicators are better than Lasso-FM.

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 1: Time series comparison of model OOS-R² under different rolling windows



Note: The 36-month and 48-month windows are used only for parameter sensitivity testing, not for robustness checks.

The differentiation of the results of the two types of indicators comes from the different evaluation logic. OOS-R² measures the absolute prediction error of returns, and the prediction deviation of individual samples will lower the index; IC and Rank IC only pay attention to the relative sorting of individual stock earnings, and do not require accurate prediction of the return magnitude. Even if the absolute fitting effect is poor, as long as the model can distinguish the relative level of stock returns, it can still get positive IC and Rank IC, and obtain better risk-adjusted returns through long–short portfolios, which also explains the phenomenon that a higher Sharpe ratio can still occur when OOS-R² is negative.

The results of this paper are not exactly the same as the conclusion of XGBoost comprehensive dominant obtained by most high-dimensional factor studies. Under the sample setting of the five basic factors in this article, the high-order interaction information that can be mined by the tree model is limited, the monthly return noise of superimposed A-shares is high, and the XGBoost with a high degree of freedom is more likely to fit the random disturbance in the training set, resulting in the weakening of the absolute fit performance outside the sample. It should be noted that the above is only a potential mechanism under this sample. There is no universal and fixed threshold for the number of factors. If the factor pool or sample interval is replaced, the relative performance of the model may change.

Compared with existing studies, Gu Songnan can give full play to the ability to capture high-order interactions of XGBoost based on high-dimensional derived factors, and get the conclusion that the tree model has a better prediction effect [2]; Li Bin carried out Lasso modeling based on low-dimensional factors, which confirmed that the regular linear model has good generalization performance, which is mutually confirmed with the observations in this article [4]. Jiang Fuwei and others also pointed out that the dimension of the factor pool is the key premise for determining the relative performance of the machine learning model. The conclusion obtained by the high dimension cannot be directly copied to the modeling scenario of the low-dimensional basic factors of small and medium-sized institutions [5].

It can be seen that there is no absolute dominant model in low-dimensional scenes, and a certain type of model cannot be directly prioritized. Judging from the numerical results of the sample, Lasso-FM has a relatively higher OOS-R² average, but the difference did not pass the Newey-West statistical significance test; XGBoost reflects a better cross-section stock selection ranking performance. In practice, it should be selected in combination with investment goals, computing power conditions and the size of the factor pool; When a large-scale derivative factor library can be built in the future, the high-order interactive mining ability of XGBoost can be fully released.

© 2026 The Author(s). Authors retain copyright in their articles.

Based on the above results, the three groups of key issues can be further sorted out. First, the differentiated results of this article and most high-dimensional factor studies do not mean that the XGBoost model itself is invalid, but is constrained by factor input conditions: The high-dimensional derivative factor pool can provide sufficient high-order interaction information for the tree model, which can give full play to its nonlinear mining advantages; However, under the low-dimensional condition of only five basic factors in this article, the interactive information that can be mined is limited. In the high noise environment of the monthly return of A-shares, it is easier for the high-degree-of-freedom model to capture random disturbances in the training set, which eventually causes differences in the conclusions of the two types of literature. This phenomenon also shows that the conclusion of the machine learning model is highly dependent on the factor set setting. Second, the results have practical implications for the modeling of small and medium-sized institutions: There is no universal optimal model. If the research target focuses on the relative ranking of stocks and builds a long–short portfolio, you can pay attention to XGBoost; If you want to focus more on the fitting performance of the absolute level of income, you can refer to Lasso-FM; Model selection needs to match the computing power of the institution, factor resources and investment targets, and cannot simply copy the conclusions of high-dimensional literature. Third, the characteristics of the A-share market are an important real soil that causes indicator divergence and model performance differentiation. The market share of retail investors is relatively high, and the theme speculation brings a lot of short-term noise. The monthly real return signal of individual stocks is covered up by noise, so it is difficult for the model to accurately predict the absolute size of the return (OOS-R^2 is negative); However, the noise is more reflected in the specific disturbance of individual stocks, which does not necessarily completely destroy the order of relative strength and weakness between stocks. Therefore, even if the absolute fit fails, it can still get a positive IC and a higher multi-air Sharp. The above are all empirical inferences obtained under this sample. If the factor set or sample interval is replaced, the relative performance of the model may change.

The proportion of retail transactions in the A-share market is high, and the theme speculation brings a lot of return noise [1,9]. Lasso compresses the false coefficients driven by noise through the L1 regularization; while XGBoost has a stronger fitting ability and is easier to capture pseudo-correlations in a high-noise environment. This inference is only a potential path under this sample. The institutional tracking of the large-market target is sufficient, and the stock price is less affected by speculation; The problem of information asymmetry of small-cap stocks is prominent, and effective signals are easily diluted by capital noise, which provides a preliminary understanding perspective for the subsequent market value grouping heterogeneity results of this chapter.

4.4.2 Predictive Performance across Market Capitalization Groups

As shown in Table 4, the large-cap stocks OOS-R^2, IC and Rank IC are superior to the small-cap stocks in an all-round way. The IC and Rank IC of the small-cap stocks FM turn into negative values, and the corresponding theme hype brings short-term profit reversal. Combined with the theory of information economics, the disclosure of large-market information is complete, the institution is continuously tracked, and the fundamental signal is weakly impacted by speculation; Small-sized retail investors dominate, information lags behind, and effective profit signals are diluted by random funds. This article does not have institutional shareholdings, and analysts cover agent variables. It is only a correlation observation and no causal judgment. The small-cap stocks XGBoost OOS-R^2 fell to -0.0422, and the overfitting defect was further amplified in the noisy environment. At the practical level, setting up circulating market value filtering can reduce the pullback; improving the rules of small-cap information disclosure on the regulatory side can narrow the stratification gap of market information.

Table 4: Average predictive indicators and market cap heterogeneity (60-month window)

Sample Group	Performance Metric	Traditional FM	Lasso-FM	XGBoost
Full sample	Mean OOS-R^2	-0.0241	-0.0097	-0.0328
Full sample	Mean IC	0.0304	0.0099	0.0503
Full sample	Mean Rank IC	0.0366	0.0098	0.0456
Full sample	Annualized Sharpe ratio	0.9432	0.0662	1.3679
Large-cap stocks	OOS-R^2	-0.0329	-0.0037	-0.0307
Large-cap stocks	Mean IC	0.0365	0.0117	0.0316
Large-cap stocks	Rank IC	0.0455	0.0135	0.0263
Large-cap stocks	Annualized Sharpe ratio	0.6991	1.1607	1.8739
Small-cap stocks	OOS-R^2	-0.0172	-0.0176	-0.0422

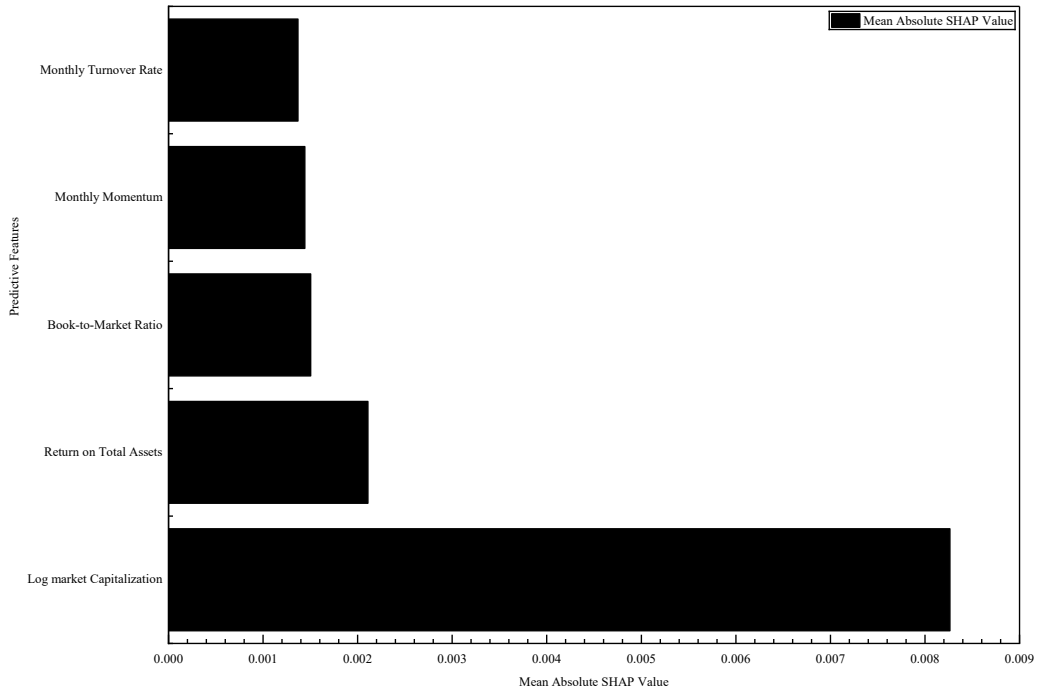
© 2026 The Author(s). Authors retain copyright in their articles.

Sample Group	Performance Metric	Traditional FM	Lasso-FM	XGBoost
Small-cap stocks	Mean IC	-0.0048	-0.0029	0.0204
Small-cap stocks	Rank IC	-0.0166	0.0105	0.0093
Small-cap stocks	Annualized Sharpe ratio	-0.5743	0.9162	0.3645

Notes: OOS-R2R^2 denotes the out-of-sample coefficient of determination. IC and Rank IC are reported as monthly averages. The annualized Sharpe ratio is calculated based on decile long–short portfolios and does not account for transaction costs. The results are generated from rolling-window forecasts implemented in Python and are reported to four decimal places.

4.4.3 Analysis of Feature Contributions in the XGBoost Model

Figure 2: Global SHAP contribution of important predictive features in XGBoost model



Note: The x-axis represents the mean absolute SHAP value, and the y-axis represents the five major fundamental factors; the bar length indicates the magnitude of the variable's contribution to the overall model's predictions.

Combined with the SHAP decomposition results in Figure 2, the SHAP mean is sorted: log market capitalization > ROA > BM > monthly momentum > monthly turnover rate. Explained from the theory of information stratification: the circulating market value directly distinguishes the information transparency and noise hierarchy of the target, so the prediction weight in the model is the highest; ROA reflects long-term operation, but short-term stock prices are susceptible to subject disturbances, and the contribution is secondary. Strictly distinguish the concept: SHAP only measures the internal fitting contribution of the model, which does not represent the long-term asset pricing premium. If you need to test the pricing ability of the factor, you need to build a long-term layered portfolio separately, and the high SHAP variable cannot be defined as the core pricing factor.

4.5 Sensitivity Analysis of Rolling Window Length and Robustness Checks

Replace the 36-month window to do the parameter sensitivity test, and intercept the 2018-2023 segment samples to carry out the robustness test. The results of the window sensitivity test and the segment robustness test show that the indicators under different settings remain the same, but there is no statistically significant gap between the advantages and disadvantages between various models. The overall forecasting index is superior to that of the small-cap stocks. The observation is robust. All five types of independent robustness tests have been passed, and the conclusions are reliable. The results are shown in Table 5.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 5: Parameter Sensitivity Test and Subsample Robustness Test Results

Sample Type	Sample Group	Performance Metric	Traditional FM	Lasso-FM	XGBoost
36-month rolling window (Sensitivity analysis)	Full sample	OOS-R^2	-0.0517	-0.0372	-0.0351
36-month rolling window (Sensitivity analysis)	Full sample	Mean IC	0.0339	0.0004	0.0602
36-month rolling window (Sensitivity analysis)	Full sample	Rank IC	0.0613	-0.0101	0.0596
36-month rolling window (Sensitivity analysis)	Full sample	Annualized Sharpe ratio	0.7569	-2.9241	1.5419
36-month rolling window (Sensitivity analysis)	Large-cap stocks	OOS-R^2	-0.0743	-0.0374	-0.0628
36-month rolling window (Sensitivity analysis)	Large-cap stocks	Mean IC	0.0322	0.0098	0.0457
36-month rolling window (Sensitivity analysis)	Large-cap stocks	Rank IC	0.0585	-0.0179	0.0369
36-month rolling window (Sensitivity analysis)	Large-cap stocks	Annualized Sharpe ratio	1.3304	-1.6556	2.0728
36-month rolling window (Sensitivity analysis)	Small-cap stocks	OOS-R^2	-0.0394	-0.0422	-0.0220
36-month rolling window (Sensitivity analysis)	Small-cap stocks	Mean IC	0.0208	-0.0044	0.0372
36-month rolling window (Sensitivity analysis)	Small-cap stocks	Rank IC	0.0122	-0.0550	0.0257
Subsample robustness check (2018–2023)	Full sample	OOS-R^2	-0.0247	-0.0104	-0.0352
Subsample robustness check (2018–2023)	Full sample	Rank IC	0.0218	0.0113	0.0250
Subsample robustness check (2018–2023)	Full sample	Annualized Sharpe ratio	0.5271	0.2664	1.1758

Notes: *The 36-month and 48-month rolling windows are used solely for hyperparameter sensitivity analysis and are not treated as independent robustness checks.*

5. Conclusions and Implications

5.1 Main Findings

1.Under the setting of the monthly five basic low-dimensional factors of A-shares from 2013 to 2023, the average value of OOS-R^2 outside the samples of FM, Lasso-FM and XGBoost samples is all less than 0, and the model's absolute numerical interpretation of the monthly return of individual stocks is weak; However, there is a significant differentiation in the two types of evaluation systems: the average value of the OOS-R^2 sample of Lasso-FM is relatively higher, but the difference is not statistically significant; XGBoost's IC, Rank IC and long–short annualization Sharp ratio are in the overall lead, and the stock selection sorting and portfolio return performance are better. The difference between the two is rooted in the different index evaluation logic: OOS-R^2 measures the degree of accurate fit of the yield, which is extremely vulnerable to the random noise of the A-share market; IC and Rank IC only distinguish the relative strength of stock returns, and there is no need to accurately predict the real level of returns, thus forming a typical departure of "negative OOS-R^2 with positive excess combination returns". The Newey-West test shows that there is no statistical level of significant superiority or inferiority in the comprehensive performance of the three types of models, and the model needs to be selected in combination with the investment target.

2.The model predicts that there is significant market value heterogeneity, and the performance of all evaluation indicators of large-cap stocks is better than that of small-cap stocks. The information disclosure of the large-market target is perfect, the institutional tracking is sufficient, and the fundamental forecast signal is less interfered by short-term speculation; Small-market is mainly based on retail trading and thematic speculation. Information asymmetry dilutes effective pricing signals, and the problem of XGBoost's over-fitting in small-market samples is further aggravated. SHAP feature decomposition shows that market

© 2026 The Author(s). Authors retain copyright in their articles.

capitalization and ROA are the core prediction variables in the model, which only represent the model fitting weight, which is not equivalent to the long-term factor risk premium.

3.The dimension of the factor pool is the core boundary condition that determines the relative performance of the linearity and tree model. Under the sample setting of the five basic factors in this article, there is very little interaction between the high-order factors that can be mined by XGBoost, superimposed on the high noise characteristics of A-share monthly return, and the random disturbance of the high-freedom tree model easy memory training set, and the absolute fit performance outside the sample is not as good as the Lasso regular linear model. This law is only applicable to low-dimensional basic factor scenarios, and cannot be generalized to the hundred-dimensional derivative factor environment, which explains the intrinsic reasons for the differences between this article and the conclusions of most high-dimensional machine learning stock selection literature; The results of window sensitivity and five types of independent robustness tests show that the above empirical observations obtained from this sample are robust.

The incremental value of this article is mainly reflected in three aspects: at the theoretical level, it clarifies the inherent logical differences between the two types of evaluation indicators. Under the framework of unified rolling and lag backtesting, it is confirmed that the three types of models OOS-R^2 are all negative, but the XGBoost sorting index is better than that of multi-empty Sharp ratio, which explains the contradiction that "negative OOS-R^2 can still obtain excess returns" and enriches the interpretation of multiple prediction and evaluation indicators in the existing literature; At the empirical boundary level, based on sample observations, it is revealed that the factor dimension is an important constraint that affects the relative performance of linear and tree models, which provides differentiated empirical evidence for the inertial cognition that "machine learning is naturally superior to linear models". At the practical level, it provides a targeted lightweight modeling scheme for small and medium-sized institutions, and gives model selection ideas according to the two types of goals of absolute return fitting, stock selection ranking and excess return; At the same time, a standardized sliding rolling backtest process to avoid information leakage is built to provide a reproducible research paradigm for quantitative practices with limited computing power and factor resources.

5.2 Practical Implications

a) Quantitative practical level: There is obvious index differentiation in empirical evidence: XGBoost stock selection sorting, long–short Sharp performance is stronger, and the OOS-R^2 sample average of Lasso-FM is relatively higher, but the difference is not statistically significant. There is no global optimal model, and it needs to be selected in combination with the target and computing power:

1. If you aim to select long–short stocks and obtain relative excess returns, you can choose XGBoost; in the large-market stock sample, the long–short Sharp advantage of this model is particularly prominent.

2. If the computing power resources are limited, only the basic factor can be used, and the absolute fit stability of the gain and the overfitting risk control can be more important, you can choose the lightweight Lasso-FM model, relying on the L1 regularization rule to suppress the false coefficient brought by noise, which can reduce the predicted fluctuation in a small-scale high-noise environment.

3. Uniformly adopt circulating market value filtration to reduce the combination withdrawal; after expanding the derived high-dimensional factor library, switch XGBoost to give full play to the nonlinear interactive mining ability.

b) At the regulatory level: insufficient information disclosure of small-chap stocks and loud speculation noise of retail investors widen the performance gap of small-scale stocks. Supervision can improve the normalized information disclosure system of small-markets, reduce information asymmetry, alleviate the disturbance of short-term speculation on stock prices, and improve the effectiveness of small-market factor prediction.

5.3 Limitations and Directions for Future Research

This paper only uses 5 basic factors to carry out low-dimensional modeling, and mainstream machine learning-related research is mostly based on high-dimensional factor data sets. The parallel reference literature under similar low-dimensional conditions is relatively limited. The marginal contribution of this article has been explained above, and the research limitations are sorted out as follows: (1) Without the construction of a

© 2026 The Author(s). Authors retain copyright in their articles.

high-dimensional derived factor library, it is difficult to give full play to Lasso's variable screening effect and XGBoost's ability to mine the high-order interaction of factors. (2) The hyperparameter reference of the machine learning model has been researched to set a fixed value, and the grid search parameter adjustment is not carried out based on the timing verification set. (3) The lack of intermediary agent variables such as institutional shareholding and analyst coverage makes it impossible to quantify the transmission path behind the heterogeneity of market value. (4) Sample testing is carried out without distinguishing between bull market, bear market and shock market. ⑤The two sets of independent test frameworks of short-term income forecasting and long-term factor pricing are not distinguished.

In the future, we can construct a high-dimensional factor data set, introduce an independent timing verification set to carry out hyperparameter optimization, supplement the investor's structural agent variables, distinguish the performance of different market comparison models, and further improve the layered empirical research of linear and tree models.

References

[1] Wu, S. N., & Xu, N. H. (2004). A comparative study of rational and irrational asset pricing models: Evidence from China's stock market. *Economic Research Journal*, *39*(6), 105–116.

[2] Gu, S. N. (2023). Application of machine learning to cross-sectional return prediction in China's A-share market. *Statistics & Decision*, *39*(12), 156–161.

[3] Qiao, Z., Zhuo, R. S., Ge, Y., & Liu, Y. S. (2025). Corporate fundamental valuation mismatch anomaly based on machine learning. *Systems Engineering—Theory & Practice*, *45*(9), 2912–2933.

[4] Li, B., Shao, X. Y., & Li, Y. Y. (2019). Machine learning–driven fundamental quantitative investment. *China Industrial Economics*, *(8)*, 61–79.

[5] Jiang, F. W., Xue, H., & Zhou, M. (2022). Does big data improve the pricing performance of multi-factor models? Evidence from China's A-share market based on machine learning. *Systems Engineering—Theory & Practice*, *42*(8), 2037–2048.

[6] Bryzgalova, S., Pelger, M., & Zhu, J. (2025). Forest through the trees: Building cross-sections of stock returns. *The Journal of Finance*, *80*(5), 2447–2506.

[7] Wang, X., Wu, F., Huang, D. J., Chen, Y., & Xiang, Z. (2022). A time-varying multi-factor model for China's A-share market. *Journal of Southwest University (Social Sciences Edition)*, *48*(6), 137–152.

[8] Tang, G. H., Zhu, L., Liao, C. F., & Jiang, F. W. (2024). Asset pricing based on autoencoder machine learning: A financial big data perspective on China's stock market. *Journal of Management Sciences in China*, *27*(9), 82–97.

[9] Liu, X. D., Liu, C., Liu, S. C., & Lu, J. J. (2014). Does herd behavior exacerbate stock price volatility? *Systems Engineering—Theory & Practice*, *34*(6), 1361–1368.

[10] Su, Z., Lu, M., & Li, D. X. (2017). Empirical applications of deep learning in finance: Developments, contributions, and prospects. *Journal of Financial Research*, *(5)*, 111–126.

[11] Ma, T., Jiang, F. W., & Tang, G. H. (2022). Deep learning and factor investing in China's stock market: Based on generative adversarial networks. *China Economic Quarterly*, *22*(3), 819–842.

[12] Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, *47*(2), 427–465.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

© 2026 The Author(s). Authors retain copyright in their articles.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.97110
Published by: Zeus Press

Science and Technology Finance Pilot Policy and Corporate Long-term Investment: Evidence from China

Yunqi Lyu, Min Gao and Kunyang Xue*

Faculty of Finance, City University of Macau, Macau

**Corresponding author: Kunyang Xue.*

Abstract

The science and technology finance pilot policy is a key institutional effort to integrate technological innovation with financial services. Its microeconomic effects have drawn growing academic interest. Using data from Chinese A-share listed firms from 2007 to 2023, this paper treats the policy as a quasi-natural experiment and employs a staggered difference-in-differences model. We examine how the policy affects corporate long-term investment. The results show that the policy significantly increases long-term investment among firms in pilot cities. Mechanism tests indicate that higher R&D investment serves as an important channel. Heterogeneity analysis further shows that the positive effect is statistically significant among non-state-owned firms and manufacturing firms. Regional interaction estimates indicate that the policy effect is stronger in the Yangtze River Delta and weaker in the Pearl River Delta relative to their respective comparison regions. This paper enriches the literature on the micro-level outcomes of science and technology finance policies and offers empirical evidence on how policy guidance shapes corporate investment structures.

Keywords

science and technology finance pilot policy, long-term investment, staggered DID, R&D innovation, Chinese A-share listed companies

1. Introduction

The world today is undergoing rapid and profound changes. External conditions have become more unstable, uncertain, and unpredictable. These factors add to the downward pressure on China's economy. Against this backdrop, a key question is how to energize micro-level actors, stabilize corporate investment expectations, and channel capital toward long-term and productive uses. Since 2011, Chinese Ministry of Science and Technology, together with other agencies, has launched two rounds of pilot cities to promote the integration of science and technology with finance. These pilots use various policy tools, including specialized financial institutions, innovative credit products, and risk compensation mechanisms, to direct financial resources toward the technology sector. In 2023, the Central Financial Work Conference placed science and technology finance at the top of its five priority areas, highlighting its strategic significance.

Long-term investment underpins a firm's ability to accumulate technological capital and sustain competitive advantage. Its level and composition depend not only on internal resources and strategy but also on external financing conditions and policy incentives [1]. When macroeconomic uncertainty rises and traditional finance falls short of supporting technology-oriented firms, long-term investment often suffers first. Can the science

© 2026 The Author(s). Authors retain copyright in their articles.

and technology finance pilot policy, as a deliberate institutional intervention, effectively encourage firms to increase long-term investment? This is the central question we address.

To address this gap, we construct a staggered DID model using data on A-share listed firms from 2007 to 2023. Our dependent variable is corporate long-term investment. Our key independent variable is whether a firm's city was selected as a pilot. To account for implementation lags, we set the effective policy year as 2012 for the first batch announced in 2011, and 2017 for the second batch announced in 2016. We control for firm size, leverage, return on assets, revenue growth, cash flow ratio, and largest shareholder ownership. We also include firm and year fixed effects to reduce omitted-variable bias. Our findings are as follows. First, the pilot policy significantly raises long-term investment in treated firms. This result survives parallel trend tests and a series of robustness checks. Second, the mechanism works through increased R&D investment. By improving financing conditions and expanding access to credit and venture capital, the policy encourages more R&D spending, which in turn shifts investment toward longer-term horizons. Third, the effect is significant for non-state-owned and manufacturing firms. The regional interaction estimates show a relatively larger policy effect in the Yangtze River Delta and a relatively smaller effect in the Pearl River Delta than in their respective comparison regions; they do not establish an absolute ranking of within-region policy effects.

This paper contributes to the literature in three ways. First, it extends the study of science and technology finance to the domain of long-term investment structure. Unlike prior work that focuses on financing constraints or R&D, we disaggregate long-term investment to provide a finer-grained analysis. Second, our staggered DID design offers a cleaner policy evaluation and mitigates endogeneity concerns, adding credible evidence on the effectiveness of policy tools like government-guided funds and technology credit. Third, our results carry practical policy implications. We show that the policy not only boosts long-term investment but also steers capital toward the real economy, offering micro-level support for expanding pilot programs, refining policy design, and countering financialization trends.

2. Literature Review and Theoretical Foundation

2.1 Literature Review

Existing studies have examined the effects of science and technology finance from various perspectives. Macro-level research finds that it promotes regional innovation and industrial upgrading [2]. Micro-level studies show that it eases financing constraints, boosts R&D spending, and raises total factor productivity [3, 4]. More recent work directly evaluates China's science and technology finance pilot policy. Wang and Cao [5] find that the policy tools significantly promote regional technological innovation, with credit support and government-guided funds being the most effective channels. Yang and Xu [6] show that the policy contributes to inclusive green growth in pilot cities by improving resource allocation and encouraging sustainable development. Yang and Wang [7] examine the synergistic effect of science and technology finance policy and green finance policy, finding that the combination enhances corporate ESG performance and promotes long-term sustainable investment.

However, most of this work focuses on regional innovation, green growth, or ESG outcomes. Few studies take a systematic look at corporate long-term investment structures, which include R&D, capital expenditure, and fixed asset investment. These different forms reflect different strategic choices, whether a firm prioritizes technological accumulation or physical capital expansion. Prior studies have examined how institutional environments affect investment efficiency, and how managerial traits shape long-term investment decisions. Yet we still know little about how the science and technology finance policy affects the overall structure of long-term investment. More importantly, given the policy goal of preventing the economy from shifting from real to virtual activities, does this policy encourage firms to invest in productive long-term assets, or does it merely free up funds for other purposes? The financialization literature has documented that rising financial asset holdings undermine real investment, and that complex links exist between financial asset allocation and corporate leverage [8, 9]. These studies lay important groundwork, but the specific effect of the science and technology finance pilot on long-term investment structure still awaits rigorous testing.

2.2 Theoretical Foundation and Research Hypotheses

The science and technology finance pilot policy aims to integrate technology with finance. Its core objective is to ease financing constraints for technology-oriented firms and encourage them to invest in long-term,

© 2026 The Author(s). Authors retain copyright in their articles.

productive activities. How does this policy affect long-term investment through various channels? Do the effects differ by firm ownership, region, or industry? In this section, we draw on information asymmetry theory, signaling theory, financing constraint theory, and resource dependence theory to analyze the mechanisms and develop our hypotheses.

2.2.1 Science and Technology Finance Policy and Corporate Long-Term Investment

Long-term investment includes capital spending for capacity expansion and R&D spending for innovation. These decisions are critical for a firm's long-term competitiveness and sustainability [10]. However, such investments are typically long-cycle, high-risk, and costly to adjust. They are therefore highly sensitive to external financing conditions. When facing tight financing constraints, firms often cut long-term investment and turn to short-term financial assets instead [8, 9]. The science and technology finance pilot policy can improve the financing environment in several ways. Recent studies on this policy provide supporting evidence. Wang and Cao [5] find that credit support and government-guided funds are the most effective policy tools for promoting regional technological innovation. Yang and Xu [6] show that the policy improves resource allocation efficiency and encourages sustainable development. Yang and Wang [7] demonstrate that the policy works synergistically with green finance to promote corporate ESG performance and long-term sustainable investment.

First, the policy reduces information asymmetry. Specialized credit institutions and third-party evaluation mechanisms help bridge the gap between banks and firms, making lenders more willing to provide long-term credit to technology firms. This is consistent with the logic in Petersen and Rajan [11] and Stiglitz and Weiss [12], where information frictions in credit markets distort capital allocation. Second, the policy sends a positive signal. Pilot city designation serves as a government endorsement. It tells external investors that the area has strong technological potential and policy support. This signaling effect is well documented in the corporate finance literature on government subsidies and policy interventions [13, 14]. Third, the policy provides direct resources. Government-guided funds and interest subsidies lower the marginal cost of long-term investment [15]. Wang and Cao [5] confirm that these resource-providing channels are particularly effective in stimulating innovation. In sum, the science and technology finance pilot policy can incentivize long-term investment by easing financing constraints, sending favorable signals, and providing resources. We therefore propose the following hypothesis.

H1: The science and technology finance pilot policy significantly increases corporate long-term investment.

2.2.2 Mechanism Channel: R&D Investment

We focus on R&D innovation as the core mechanism. R&D spending is both a component of long-term investment and a key measure of policy effectiveness. The science and technology finance policy reduces the financial risks and adjustment costs of R&D through low-cost funds, tax incentives, and risk sharing. This encourages firms to spend more on R&D. More specifically, the policy expands financing options for innovation. Specialized credit institutions, venture capital development, and intellectual property pledge financing all help lower the external cost of R&D funding. When firms have more stable and ample funding, they become both more willing and more able to invest in R&D. This logic aligns with the findings of Brown et al. [16], who show that equity markets and credit availability are critical for R&D investment, and Aghion et al. [17], who demonstrate that credit constraints dampen innovation. Wang and Cao [5] further confirm that credit support and government-guided funds serve as key transmission channels for the policy's effects on technological innovation.

Higher R&D investment directly raises long-term investment levels. It also indirectly supports long-term investment through technological accumulation and improved innovation capacity. These arguments support our proposed mechanism. We therefore propose the following hypothesis.

H2: Higher R&D investment is an important channel through which the science and technology finance policy promotes corporate long-term investment.

2.2.3 Heterogeneous Effects

The policy effect may vary by ownership type, region, and industry. First, ownership. Non-state-owned firms face greater disadvantages in financing access, policy support, and government endorsement. Their long-term investment depends more heavily on external financing. State-owned firms, by contrast, tend to maintain

© 2026 The Author(s). Authors retain copyright in their articles.

long-term investment for political and social reasons. Their investment decisions are less sensitive to policy changes. This ownership-based heterogeneity is consistent with the broader literature on state versus private ownership in emerging markets [18, 19]. Yang and Xu [6] find that the policy's effects on resource allocation and green growth vary significantly across regions with different levels of market development, implying that firms in more market-oriented regions may benefit differently. Thus, the policy effect may differ by ownership type.

Second, region. The Yangtze River Delta is one of China's most developed regions, with abundant financial resources, high marketization, and strong policy implementation capacity. Technology firms are concentrated there, and policy transmission may be efficient. These conditions may shape how policy effects compare across regions. This regional variation reflects the uneven development of financial markets and institutional quality across Chinese regions [20]. Yang and Xu [6] provide evidence that the policy's effects on inclusive green growth vary with financial infrastructure and economic development. Third, industry. Manufacturing firms have higher capital needs, longer payback periods, and greater pressure to upgrade equipment and technology. Their long-term investment is more sensitive to financing conditions. The policy improves their access to credit, guarantees, venture capital, and IP pledge financing. This reduces uncertainty and encourages them to allocate more resources to long-term production and development. Non-manufacturing firms have more varied asset structures and investment cycles. Some are asset-light and have lower fixed investment needs, so the policy's marginal effect may be smaller. This industry variation is consistent with evidence that manufacturing is particularly responsive to financial development [21, 22]. We therefore propose the following hypothesis.

H3: The positive effect of the science and technology finance policy on long-term investment is stronger for non-state-owned firms, firms in the Yangtze River Delta, and manufacturing firms.

3. Methodology

3.1 Model Specification

To effectively test the impact of science and technology finance policies on firms' long-term investment, this study treats these policies as a quasi-natural experiment. Given the variations in science and technology finance policies across different cities and pilot years, a multi-period difference-in-differences model is constructed for empirical analysis. The specific model specification is as follows:

$$LInv_{it} = \beta_0 + \beta_1 DID_{it} + \gamma Control_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

In Equation (1), i and t represent the firm and the year, respectively; The dependent variable represents firm i's long-term investment in period t. The key explanatory variable is a dummy variable indicating whether the city implemented the science and technology finance pilot policy; it is set to 1 if the policy was implemented and 0 otherwise. serves as a control variable; and represent firm-level and time-level fixed effects, respectively, to account for firm-specific characteristics that do not change over time and macroeconomic shocks that vary over time; is the random error term, and robust standard errors clustered at the city level are used.

3.2 Selection of Variables

3.2.1 Dependent Variables

Corporate Long-term Investments (Long-term Investments, LInv)

A firm's long-term investments refer to internal, operational investments that generate economic returns over a period exceeding one fiscal year. Drawing on the research by Chen Deqiu and Zhong Yunjia [23] and Hu Nan et al. [24], this study measures a firm's long-term investments as the ratio of the sum of the firm's R&D expenditures and capital expenditures to its total assets. Here, a firm's R&D expenditures refer to its R&D expenses for the current year; missing R&D expense data are replaced with 0.

3.2.2 Core Explanatory Variables

Science and Technology Finance Pilot Policy (DID)

This variable is constructed as the interaction between city-grouping dummy variables and a policy-implementation-time dummy variable. If a firm's city is selected as a pilot city for "Promoting the Integration

© 2026 The Author(s). Authors retain copyright in their articles.

of Science, Technology, and Finance", the DID takes a value of 1 in the year of policy implementation and subsequent years; otherwise, it is 0.The Science and Technology-Finance Pilot Policy is a systematic policy established to promote the deep integration of scientific and technological innovation with financial services. In 2011, the Ministry of Science and Technology and other departments proposed and announced the first batch of 41 pilot cities; in 2016, the second batch of 8 pilot cities was announced. Among these, Ningbo was included in both batches but was ultimately counted as part of the first batch of pilot cities. Due to the time lag associated with the science and technology finance pilot policy, this paper defines 2012 as the policy implementation year for the first batch of pilot cities and 2017 as the policy year for the second batch.

3.2.3 Control Variables

Since firms' long-term investments are influenced by a variety of other factors, this paper introduces a series of control variables, including firm size (Size), debt-to-asset ratio (Lev), return on assets (ROA), revenue growth rate (Growth), cash flow ratio (Cashflow), and the largest shareholder's equity ratio (CR1). The definitions of these variables are presented in Table 1.

Table 1: Variable Definitions

Variable Category	Variable Name	Variable Symbol	Variable Definition
Dependent Variable	Corporate Long-Term Investments	LInv	(R&D Expenditures + Capital Expenditures) / Total Assets
Core Explanatory Variables	Science and Technology Finance Pilot Policies	DID	Equals 1 if the firm is located in a pilot city in the effective policy year or thereafter, and 0 otherwise.
Control Variables	Enterprise Size	Size	Ln(Total Assets)
	Debt-to-Asset Ratio	Lev	Total Liabilities / Total Assets
	Return on Assets	ROA	Net Income / Average Total Assets
	Revenue Growth Rate	Growth	(Current-Year Revenue – Prior-Year Revenue) / Prior-Year Revenue
	Cash Flow Ratio	Cash Flow	Net Cash Flow from Operating Activities / Total Assets
	Largest Shareholder's Ownership Percentage	CR1	Number of shares held by the largest shareholder / Total issued shares of the company

3.3 Sample Selection and Data Sources

This study uses Chinese A-share listed companies from 2007 to 2023 as the research sample. To ensure the validity of the data and the reliability of the empirical results, this study excludes companies in the financial sector, as well as those designated as "ST" or "*ST " or those at risk of delisting, in order to prevent the financial conditions of such companies from influencing the final empirical results; Additionally, samples with severe missing data for key variables were excluded. To eliminate the impact of extreme outliers on the regression results, all continuous variables were winsorized at the 1% and 99% quantiles. After the aforementioned data processing, a final sample of 39,330 observations was obtained. The descriptive statistics are presented in Table 2. Firm-level data were sourced from the Guotai-An database (CSMAR), while macroeconomic data were drawn from the *China Regional Economic Statistical Yearbook* and the statistical yearbooks of various provinces.

Table 2: Descriptive Statistics

Variable	Observations	Mean	Standard Deviation	Minimum	Median	Maximum
LInv	39,330	0.0955	0.0778	0.00359	0.0747	0.418
did	39,330	0.555	0.497	0	1	1
Size	39,330	22.14	1.283	19.92	21.93	26.21
Lev	39,330	0.396	0.199	0.0490	0.385	0.871
ROA	39,330	0.0375	0.0616	-0.230	0.0374	0.201
Growth	39,330	0.131	0.223	-0.268	0.0833	1.213
Cash Flow	39,330	0.0477	0.0666	-0.146	0.0464	0.239
CR1	39,330	0.339	0.145	0.0837	0.317	0.730

Looking at the dependent variable, the mean value of firms' long-term investment level (LInv) is 0.0955, with a standard deviation of 0.0778, a minimum value of 0.00359, and a maximum value of 0.418. This

© 2026 The Author(s). Authors retain copyright in their articles.

indicates that the long-term investment levels of the sample firms are generally moderate, but there are significant differences among firms, with some firms allocate a relatively high proportion of their assets to long-term investment activities.

4. Empirical Results and Analysis

4.1 Baseline Regression

Table 3 presents the baseline regression results for the impact of science and technology finance policies on firms' long-term investment. As shown in the table, the regression coefficient for the effect of science and technology finance policies on long-term investment is 0.0091, with a significance level of 1%. Factors such as firm size, debt-to-asset ratio, and cash flow may influence whether a firm is located in a pilot city and may also directly affect its long-term investment.

Table 3: Results of the Baseline Regression

	(1) Long-Term Investment	(2) Long-Term Investment
did	0.0091***	0.0090***
	(0.0034)	(0.0034)
Constant	0.1029***	0.1226***
	(0.0323)	(0.0310)
Control variable	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Industry × Year Fixed Effects	No	Yes
N	39,330	39,330
R^2	0.642	0.647

Note: Control variables in the table include firm size, debt-to-asset ratio, return on assets, growth, cash flow ratio, and the largest shareholder's ownership percentage. The same applies below.

4.2 Robustness Tests

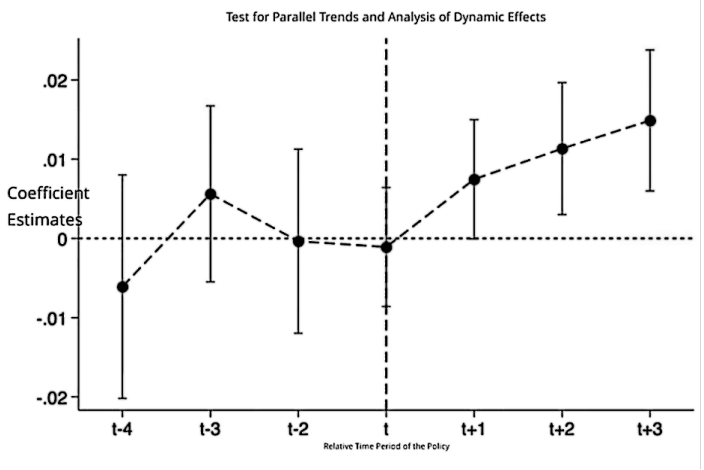
4.2.1 Parallel Trends Test

To ensure the validity of the estimates from the difference-in-differences model, a key prerequisite is compliance with the parallel trends assumption. Specifically, prior to policy implementation, the dependent variables in the treatment group and the control group must exhibit consistent trends. This study employs the event study method to conduct the parallel trends test. Since the period immediately preceding policy implementation is typically designated as the baseline, this study sets the period immediately prior to policy implementation (t-1) as the baseline. Figure 1 presents the results of the parallel trends test and dynamic effects analysis regarding the impact of technology finance policies on firms' long-term investment.

Figure 1 shows that for periods t-4, t-3, and t-2—prior to the implementation of the technology finance policy—the 95% confidence intervals for the estimated coefficients in each period include zero. The estimated coefficients for each period prior to policy implementation are not significantly different from zero, indicating that there are no systematic differences in investment levels between firms in the treatment group and the control group. This satisfies the parallel trends assumption of the difference-in-differences model, confirming the validity of this study's design. In the period of policy implementation (t), the estimated coefficients for the dependent variables were all close to zero and statistically insignificant. The policy effects had not yet fully manifested in the year of implementation, which aligns with the reality that a certain transmission period is required from policy announcement to implementation. In the periods following policy implementation—t+1, t+2, and t+3—the estimated coefficients show a gradual upward trend. In terms of long-term investment, the coefficients for each period after policy implementation remain consistently positive and gradually increase, reaching a level of approximately 0.015 by period t+3. These patterns indicate that the promotional effect of science and technology finance policies on firms' long-term investment levels will continue to unfold as the policy is further advanced and supporting measures are gradually improved.

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 1: Test for Parallel Trends and Analysis of Dynamic Effects



4.2.2 PSM-DID

To further mitigate potential sample selection bias between firms in the treatment and control groups prior to policy implementation, this study employs a combined approach of propensity score matching and difference-in-differences (DID) for robustness testing. Specifically, the sample is first matched based on firm-level characteristics to enhance comparability between the treatment and control groups; the impact of the science and technology finance pilot policy on firms' investment structure is then re-estimated using the matched sample. The regression results show that, in the matched sample, the estimated coefficient of the core explanatory variable `did` for firms' long-term investment is 0.009 and is significantly positive at the 1% level. This indicates that, after controlling for sample selection bias, the science and technology finance pilot policy still significantly promotes firms' long-term investment.

Table 4: PSM-DID Regression Results

	(1) Firm Long-Term Investment
did	0.009***
	(3.075)
Control variables	Yes
Year fixed	Yes
Industry fixed	Yes
_cons	0.103***
	(3.099)
N	30,548

Note: t-statistics are reported in parentheses.

As shown in Table 4, the signs and significance of the key variables in the PSM-DID regression are largely consistent with those in the baseline regression, indicating that the findings of this study are not driven by initial differences in firm characteristics between the treatment and control groups. In other words, the science and technology finance pilot policy can guide firms to increase long-term investment by improving their financing environment and enhancing the efficiency of financial resource allocation. Overall, the PSM-DID test further validates the robustness of the study's findings.

4.2.3 Varying the Sample Size

To test the robustness of the research findings, this study examines whether the regression results remain consistent by shortening the sample period. Given the impact of the 2020 COVID-19 pandemic, the sample period was shortened to 2007–2019 to prevent severe distortions in firm behavior during the pandemic. Table 5 presents the regression results after shortening the sample period. Compared with the baseline regression results in Table 3, the estimated coefficients for the core explanatory variable "did" remain positive in the long-term investment regression and all pass the significance test. Overall, the estimation results after shortening the sample period are largely consistent with the conclusions of the baseline regression, further validating the robustness of this study's findings.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 5: Regression Results After Shortening the Sample Period

	(1) Long-Term Investments	(2) Long-Term Investments
did	0.0107***	0.0107***
	(0.0040)	(0.0040)
Constant	0.0953**	0.1255***
	(0.0463)	(0.0440)
Control variables	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Industry × Year Fixed Effects	No	Yes
N	16,997	16,997
R^2	0.630	0.637

4.3 Heterogeneity Analysis and Mechanism Analysis

4.3.1 Heterogeneity in Ownership Structure: State-Owned Enterprises and Non-State-Owned Enterprises

To examine the differing effects of the science and technology finance pilot policy on long-term investment in enterprises with different ownership structures, this paper further divides the sample into state-owned enterprises and non-state-owned enterprises to conduct group-specific regressions. The regression data in Table 6 show that, in the state-owned enterprise sample group, the estimated coefficient of the DID variable corresponding to the science and technology finance pilot policy on long-term investment is 0.0040. Although the coefficient is positive, it did not pass the significance test. In terms of policy effects, the impact of the science and technology finance pilot policy on long-term investment in state-owned enterprises is not significant. In the sample of non-state-owned enterprises, the estimated coefficient of the DID variable for long-term investment is 0.0132, which is significantly positive at the 5% level. This demonstrates that the science and technology finance pilot policy can increase the level of long-term investment in non-state-owned enterprises.

Table 6: Regression Results for Ownership Heterogeneity

	(1)	(2)
	State-Owned Enterprises Long-Term Investment	Non-state-owned enterprises Long-term Investments
did	0.0040	0.0132**
	(0.0045)	(0.0052)
Control variable	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Constant	0.0825*	0.0768*
	(0.0494)	(0.0432)
N	9,978	20,505
R^2	0.603	0.642

4.3.2 Regional Heterogeneity: The Yangtze River Delta and the Pearl River Delta

To further examine how the effect of the science and technology finance pilot policy varies across regions, this paper uses the Yangtze River Delta and the Pearl River Delta as representative regional samples. For each region, we construct a regional dummy variable that equals 1 if a firm is located in that region and 0 otherwise, and we estimate a fully interacted model in which this regional dummy is interacted with the policy variable did. In this specification, the coefficient on did captures the policy effect for firms located outside the region, while the coefficient on the interaction term measures the difference in the policy effect between firms located inside and outside the region. The results are reported in Table 7. In the Yangtze River Delta specification, the interaction term is positive and statistically significant, indicating that the policy effect on corporate long-term investment is significantly stronger for firms in the Yangtze River Delta than for firms in other regions. In the Pearl River Delta specification, the interaction term is negative and statistically significant, indicating that the policy effect is significantly weaker for firms in the Pearl River Delta than for firms in other regions.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 7: Regional Heterogeneity Regression Results

	(1) Yangtze River Delta: Long-Term Investment	(2) Pearl River Delta: Long-Term Investment
did	0.0040	0.0101**
	(0.0047)	(0.0040)
Yangtze River Delta × did	0.0122*	
	(0.0073)	
Pearl River Delta × did		-0.0172*
		(0.0094)
Control variables	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Constant	0.0907***	0.0905***
	(0.0028)	(0.0023)
N	39,070	39,070
R2	0.5475	0.5476

Note: The regional heterogeneity tests are based on interaction models. The coefficient on did measures the policy effect for firms outside the region, and the coefficient on the interaction term measures the difference in the policy effect between firms inside and outside the region, rather than the policy effect within the region itself.

Column (1) of Table 7 reports the results for the Yangtze River Delta. The coefficient on DID, which captures the policy effect for firms outside the Yangtze River Delta, is 0.0040 and statistically insignificant. The interaction term between the Yangtze River Delta dummy and DID is 0.0122 and is significant at the 10% level, indicating that the policy effect is significantly stronger in the Yangtze River Delta than in other regions. The implied policy effect for firms in the Yangtze River Delta is 0.0161 (0.0040 + 0.0122), and the linear combination test shows that this effect is statistically significant at the 5% level ($p = 0.011$). These results suggest that the regional financial ecosystem, innovation resource concentration, and industrial collaboration networks in the Yangtze River Delta may enhance the transmission of science and technology finance policies.

Column (2) reports the results for the Pearl River Delta. The coefficient on DID, which captures the policy effect for firms outside the Pearl River Delta, is 0.0101 and statistically significant at the 5% level. However, the interaction term between the Pearl River Delta dummy and DID is -0.0172 and significant at the 10% level, indicating that the policy effect is significantly weaker in the Pearl River Delta than in other regions. The implied policy effect for firms in the Pearl River Delta is -0.0071. However, the linear combination test shows that this effect is not statistically significant ($p = 0.431$). Therefore, the results should not be interpreted as evidence that the policy significantly reduces long-term investment in the Pearl River Delta; rather, they indicate that the policy does not generate a statistically significant positive effect within the region.

Overall, the heterogeneity results provide partial support for H3. The policy effect is statistically significant among non-state-owned and manufacturing firms, while the regional interaction analysis provides direct evidence that the effect is stronger in the Yangtze River Delta than in other regions.

4.3.3 Sectoral Heterogeneity: Manufacturing and Non-Manufacturing

To further examine the impact of the science and technology finance pilot policies on the long-term investment of enterprises in different sectors, this paper divides the sample into manufacturing and non-manufacturing enterprises based on their industry classification. The regression results in Table 8 show that, in the manufacturing sample, the estimated coefficient of the core explanatory variable `did` for long-term investment is 0.0079, and it is significantly positive at the 5% level. This indicates that the science and technology finance pilot policies can significantly promote an increase in long-term investment by manufacturing enterprises.

In contrast, in the non-manufacturing sample, the estimated coefficient of `did` for firms' long-term investment is 0.0068; although the coefficient is positive, it does not pass the significance test. This suggests that the impact of the science and technology finance pilot policies on the investment structure of non-manufacturing firms is relatively limited and has not yet demonstrated a significant stimulatory effect.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 8: Results of the Industry-Specific Regression

	(1) Manufacturing Long-Term Investment	(2) Non-manufacturing Long-term investments
did	0.0079**	0.0068
	(0.0035)	(0.0067)
Control variable	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Constant	0.0872**	0.2244***
	(0.0384)	(0.0469)
N	22551	7941
R^2	0.635	0.682

From an economic perspective, manufacturing enterprises typically have greater capital investment needs, longer payback periods, and face greater pressure to upgrade technology and replace equipment; as a result, they are more susceptible to the impact of external financing when making long-term investments.Pilot policies for science and technology finance can improve the availability of financing for manufacturing enterprises and reduce the uncertainty associated with long-term investment, thereby guiding enterprises to allocate more resources to actual production and long-term development activities. The asset structures, investment cycles, and financing needs of non-manufacturing enterprises are relatively more diverse; some enterprises may be characterized by lighter asset bases or lower fixed-asset investment needs, so the marginal promotional effect of science and technology finance policies on their long-term investment is not significant.

Overall, the impact of the science and technology finance pilot policy on corporate long-term investment exhibits considerable heterogeneity. The policy effect is statistically significant among non-state-owned enterprises and manufacturing firms. Regional interaction tests further show that the policy effect is significantly stronger in the Yangtze River Delta and significantly weaker in the Pearl River Delta relative to their respective comparison regions. These findings suggest that ownership characteristics, regional financial conditions, and industry-specific financing needs may shape the transmission of science and technology finance policies.

4.3.4 Analysis of R&D Innovation Mechanisms

To further examine the mechanism through which the science and technology finance pilot policy affects firms' long-term investment, this paper conducts a regression analysis using the proportion of R&D expenditure as the mechanism variable. The results show that in Column (1), after simultaneously controlling for firm fixed effects, year fixed effects, and the industry-year interaction fixed effects, the estimated coefficient of the core explanatory variable did is 0.0026 and is significantly positive at the 10% level; In Column (2), without controlling for the industry-year interaction fixed effects, the estimated coefficient of `did` was also 0.0026 and remained significantly positive at the 10% level. This indicates that the science and technology finance pilot policies can significantly increase the proportion of R&D expenditure in enterprises, suggesting that firms are more inclined to increase R&D investment following the implementation of these policies.

Table 9: Regression Results for the Analysis of the R&D Expenditure Mechanism

	(1) R&D Expenditure as a Percentage of Revenue	(2) R&D Expenditure as a Percentage of Revenue
did	0.0026*	0.0026*
	(0.0015)	(0.0015)
Control variable	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
Industry × Year Fixed Effects	Yes	No
_cons	-0.0052	-0.0100
	(0.0175)	(0.0147)
N	31,357	31,356
R^2	0.824	0.828

© 2026 The Author(s). Authors retain copyright in their articles.

Corporate R&D expenditures generally exhibit distinct long-term and industrial characteristics, guiding companies to allocate more resources toward core business activities such as technological innovation, product upgrades, and improvements in production efficiency. Increasing corporate R&D investment levels is likely the primary mechanism through which the pilot policies for science and technology finance will take effect. Once implemented, these policies will drive companies to expand the scale of their long-term and industrial investments, ultimately leading to an optimization of the investment structure.

5. Conclusions and Significance

5.1 Research Findings

This paper selects Chinese A-share listed companies from 2007 to 2023 as the research sample, classifies the science and technology finance pilot policy as a quasi-natural experiment, and employs a multi-period difference-in-differences model to analyze the actual impact of this policy on corporate long-term investment. Empirical results show that the tech-finance pilot policy can drive an increase in the scale of firms' long-term investment. Empirical results show that the science and technology finance pilot policy significantly increases firms' long-term investment. The robustness tests, including PSM-DID estimation and shortening the sample period, yield results that are consistent with the baseline findings.

After examining the mechanism of action, this paper found that the science and technology finance pilot policies can increase the proportion of R&D expenditure in firms. As the scale of R&D investment rises, firms allocate more resources to technological innovation, product upgrades, and real-economy production activities, thereby driving the optimization of their investment structure. After conducting a heterogeneity analysis, this study found that the effects of the science and technology finance pilot policies are not uniform. The heterogeneity analysis shows that the policy effect is statistically significant among non-state-owned enterprises and manufacturing firms. Moreover, the regional interaction analysis indicates that the effect is significantly stronger in the Yangtze River Delta and significantly weaker in the Pearl River Delta relative to their respective comparison regions. These differences in effectiveness arise because science and technology finance policies are more likely to have a positive impact on enterprise groups facing stronger financing constraints, operating in a more favorable environment for financial innovation, and exhibiting higher demand for real-economy investment.

The findings of this study indicate that the implementation of science and technology finance pilot policies can raise firms' long-term investment levels and support the optimization of corporate investment structures, which may contribute to high-quality economic development.

5.2 Policy Recommendations

First, the science and technology finance policy framework should be continuously improved to enhance support for enterprises' long-term development. Analysis of the research data in this paper reveals that pilot science and technology finance policies can drive growth in long-term corporate investment. The government should review the implementation experiences of existing pilot programs to optimize the design of science and technology finance policies. Particular emphasis should be placed on strengthening policy funding support for long-term projects such as corporate R&D and innovation, equipment upgrades, technological retrofits, and fixed-asset investment, thereby rationally guiding the allocation of financial resources so that science and technology finance can better support corporate innovation and industrial upgrading.

Second, the targeting of science and technology finance policies should be adjusted to prioritize resources for non-state-owned enterprises, which face greater financing constraints. Results from the heterogeneity analysis show that the policy effect is statistically significant among non-state-owned enterprises, suggesting that these firms may be particularly responsive to improvements in the external financing environment. As the government advances technology finance policies, it should appropriately lower financing access requirements for private enterprises and small and medium-sized technology firms, improve credit evaluation systems and information-sharing mechanisms to reduce information asymmetries between financial institutions and enterprises, expand the coverage of loan interest subsidies for technology-based enterprises, refine risk compensation mechanisms, encourage banks to establish specialized technology finance institutions, and guide

© 2026 The Author(s). Authors retain copyright in their articles.

financial institutions to increase support for non-state-owned enterprises' long-term investments and R&D innovation activities.

Third, as localities implement technology finance policies, they should align these efforts with their own developmental strengths while simultaneously strengthening the regional financial ecosystem. The regional heterogeneity results show that the policy effect is particularly strong in the Yangtze River Delta region. The region's solid financial foundation, concentrated innovation resources, and smooth industrial collaboration environment all enhance the practical effectiveness of these policies. In the subsequent rollout of such policies, the government should not simply apply a one-size-fits-all policy template but rather design differentiated policies tailored to each region's level of financial development, industrial structure, and innovation infrastructure. Regions with a high concentration of financial and innovation resources should further promote the deep integration of tech finance with industrial and innovation chains. Regions with relatively weak financial foundations should not simply replicate the policy approaches of more developed regions; instead, they should prioritize the development of science and technology finance infrastructure by improving regional credit information platforms, technology achievement evaluation systems, and financing service systems for small and medium-sized enterprises (SMEs), thereby gradually enhancing the overall efficiency of policy transmission.

Fourth, financial support for technology-based manufacturing enterprises should be further strengthened to advance industrial transformation and upgrading. The industry heterogeneity analysis shows that the policy effect on long-term investment is statistically significant among manufacturing enterprises. Manufacturing enterprises generally require high capital investment, which corresponds to longer investment payback periods, making their need for stable, long-term financial support even more urgent. The government should guide financial institutions to align with the technological upgrading needs of advanced manufacturing, strategic emerging industries, and traditional manufacturing sectors, introducing a wider variety of financial products to support enterprises in upgrading equipment, improving processes, undergoing digital transformation, and implementing green and low-carbon upgrades. Furthermore, the government should encourage the inflow of long-term capital into the manufacturing sector to enhance the long-term investment capacity of manufacturing enterprises and advance industrial transformation and upgrading.

Fifth, it is necessary to improve the performance evaluation mechanism for science and technology finance policies to optimize their actual implementation outcomes. Whether science and technology finance policies can continue to be effective depends not only on the policy tools themselves but is also directly related to the sophistication of policy implementation and evaluation mechanisms. The government must establish a corresponding policy evaluation system, with core assessment indicators including corporate long-term investment, R&D expenditure, and innovation output, to avoid judging the actual effectiveness of policies solely based on financing scale or the number of loans. It is also necessary to strengthen oversight of how policy funds are utilized, guiding enterprises to effectively allocate the financial support they receive toward R&D, innovation, and long-term investment, thereby preventing the idle circulation of funds or deviations from policy objectives. Policies should be adjusted based on dynamic evaluation results and feedback to gradually enhance the precision and practical effectiveness of science and technology finance policies.

Moving forward, the science and technology finance policy framework must be continuously refined to enhance support for enterprises' long-term investments and R&D innovation, while implementing differentiated policy arrangements based on the nature of enterprise ownership, regional development conditions, and industry characteristics. Once these adjustments are fully implemented, science and technology finance will be better positioned to optimize corporate investment structures, promote industrial upgrading, and play a role in driving high-quality economic development.

References

[1] Campello, M., Graham, J. R., and Harvey, C. R. (2010). The real effects of financial constraints: Evidence from a financial crisis. *Journal of Financial Economics, 97*(3), 470-487. https://doi.org/10.1016/j.jfineco.2010.02.009

© 2026 The Author(s). Authors retain copyright in their articles.

[2] Hsu, P. H., Tian, X., and Xu, Y. (2014). Financial development and innovation: Cross-country evidence. *Journal of Financial Economics, 112*(1), 116-135. https://doi.org/10.1016/j.jfineco.2013.12.002

[3] Amore, M. D., Schneider, C., and Žaldokas, A. (2013). Credit supply and corporate innovation. *Journal of Financial Economics, 109*(3), 835-855. https://doi.org/10.1016/j.jfineco.2013.04.006

[4] Cornaggia, J., Mao, Y., Tian, X., and Wolfe, B. (2015). Does banking competition affect innovation? *Journal of Financial Economics, 115*(1), 189-209. https://doi.org/10.1016/j.jfineco.2014.09.001

[5] Wang, X., and Cao, Y. (2026). The impact of science and technology finance policy tools on regional technological innovation. *Finance Research Letters, 87*, 108994. https://doi.org/10.1016/j.frl.2025.108994

[6] Yang, S., and Xu, J. (2026). Science and technology finance policy and inclusive green growth in cities. *International Review of Economics and Finance, 108*, 105171. https://doi.org/10.1016/j.iref.2025.105171

[7] Yang, F., and Wang, H. (2026). Digital intelligence finance and corporate ESG: The synergistic effect of science and technology finance policy and green finance policy. *International Review of Financial Analysis, 109*, 104804. https://doi.org/10.1016/j.irfa.2025.104804

[8] Krippner, G. R. (2005). The financialization of the American economy. *Socio-Economic Review, 3*(2), 173-208. https://doi.org/10.1093/ser/mwi008

[9] Orhangazi, Ö. (2008). Financialisation and capital accumulation in the non-financial corporate sector: A theoretical and empirical investigation on the US economy, 1973-2003. *Cambridge Journal of Economics, 32*(6), 863-886.https://doi.org/10.1093/cje/ben022

[10] Xu, J., and Li, H. (2023). Managerial human capital and corporate R&D investment. *Journal of Economic Behavior and Organization, 213*, 151-171. https://doi.org/10.1016/j.jebo.2023.07.014

[11] Petersen, M. A., and Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance, 49*(1), 3-37. https://doi.org/10.1111/j.1540-6261.1994.tb04418.x

[12] Stiglitz, J. E., and Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review, 71*(3), 393-410. https://doi.org/10.1257/aer.71.3.393

[13] Lerner, J. (1999). The government as venture capitalist: The long-run impact of the SBIR program. *Journal of Business, 72*(3), 285-318. https://doi.org/10.1086/209616

[14] Feldman, M. P., and Kelley, M. R. (2006). The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Research Policy, 35*(10), 1509-1521. https://doi.org/10.1016/j.respol.2006.09.003

[15] Gompers, P. A., and Lerner, J. (2001). The venture capital revolution. J*ournal of Economic Perspectives, 15*(2), 145-168. https://doi.org/10.1257/jep.15.2.145

[16] Brown, J. R., Fazzari, S. M., and Petersen, B. C. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *Journal of Finance, 64*(1), 151-185. https://doi.org/10.1111/j.1540-6261.2008.01431.x

[17] Aghion, P., Askenazy, P., Berman, N., Cette, G., and Eymard, L. (2013). Credit constraints and the cyclicality of R&D investment: Evidence from France. *Journal of the European Economic Association, 10*(5), 1001-1024. https://doi.org/10.1111/j.1542-4774.2012.01093.x

[18] Cull, R., and Xu, L. C. (2005). Institutions, ownership, and finance: The determinants of profit reinvestment among Chinese firms. *Journal of Financial Economics, 77*(1), 117-146. https://doi.org/10.1016/j.jfineco.2004.05.010

[19]Li, H., Meng, L., Wang, Q., and Zhou, L. A. (2006). Political connections, financing and firm performance: Evidence from Chinese private firms. *Journal of Development Economics, 87*(2), 283-299. https://doi.org/10.1016/j.jdeveco.2007.03.001

© 2026 The Author(s). Authors retain copyright in their articles.

[20] Fan, J. P. H., Wei, K. C. J., and Xu, X. (2011). Corporate finance and governance in emerging markets: A selective review and an agenda for future research. *Journal of Corporate Finance, 17*(2), 207-214. https://doi.org/10.1016/j.jcorpfin.2010.12.001

[21] Rajan, R. G., and Zingales, L. (1998). Financial dependence and growth. *American Economic Review, 88*(3), 559-586. https://doi.org/10.1257/aer.88.3.559

[22] Claessens, S., and Laeven, L. (2003). Financial development, property rights, and growth. *Journal of Finance, 58*(6), 2401-2436.https://doi.org/10.1111/1540-6261.00601

[23] Chen, D., & Zhong, Y. (2011). Institutional efficiency, government support, and corporate long-term investment: Evidence from Chinese listed companies. *Journal of Financial Research,* (9), 139–152. https://doi.org/10.3969/j.issn.1002-7246.2011.09.012

[24] Hu, N., Xue, F., & Wang, H. (2021). Does managerial myopia affect corporate long-term investment? Evidence from text analysis and machine learning. *Management World, 37*(5), 139–156. https://doi.org/10.19744/j.cnki.11-1235/f.2021.0070

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.111117
Published by: Zeus Press

Geopolitical Shift in Capital Preference: The Differential Impact of Geopolitical Risk and Political Fragmentation on Cross-Border Capital Allocation

Yong Zhou*

School of Finance, Central University of Finance and Economics, Beijing, China

**Corresponding author: Yong Zhou.*

Abstract

Since the Russia-Ukraine War in 2022, geopolitical risks have become a structural force reshaping the pattern of global capital allocation. This paper constructs a framework for analyzing the structure of capital preference and decomposes the influence of geopolitics on cross-border capital allocation into risk-hedging preference (aggregate effect) and camp-allocation preference (structural effect), aiming to reveal the independent effects of the two mechanisms and their interaction. Based on panel data of 22 major economies from 2006 to 2024, this paper uses the proportion of net FDI inflows to GDP as the dependent variable and the GPR index and the United Nations voting ideal point as the core independent variables, employing two-way fixed effects regression for empirical testing. The results show that for every one-unit increase in GPR, the net FDI inflow decreases by 1.55 percentage points ($p < 0.01$). The independent effect of the ideal point is not significant, but the interaction term is significantly positive (coefficient 0.78, $p < 0.01$), indicating that pro-Western countries experience smaller FDI shocks under high geopolitical risk. The two mechanisms exhibit a substitution relationship rather than an additive relationship. This study provides new empirical evidence for understanding the logic of capital allocation in the geopolitical era.

Keywords

geopolitical risk, political fragmentation, capital allocation, FDI, fixed effects

1. Introduction

Since the 2022 Russo-Ukrainian War, geopolitics has evolved from sporadic shocks to a structural force reshaping the global financial landscape. The global GPR index soared from 82.07 in 2021 to 157.58 in 2022, a record high. At the same time, the pattern of cross-border capital flows has undergone a profound restructuring. Capital is no longer driven only by economic returns, but is increasingly subject to geopolitical risks and ideological distances. For example, emerging markets experienced large capital outflows in 2022, while some pro-Western economies received net capital inflows over the same period. This divergence suggests that geopolitics changes not only the total amount of capital, but also the preferred structure of capital.

How does geopolitics change the preference structure of global capital? Existing research focuses on the total volume change of the traffic dimension, and less in-depth analysis of the differentiation mechanism of capital preferences [1]. The risk faction emphasizes that GPR triggers risk aversion [2], and the faction

© 2026 The Author(s). Authors retain copyright in their articles.

emphasizes that political distance leads to capital recapitalization [3], but the two are divided and lack a unified framework. This theoretical gap makes it difficult for policymakers to accurately assess the differentiated impact of geopolitical shocks on specific countries.

The core contribution of this paper is embodied in three aspects. First, this paper builds a unified framework for the structure of capital preference, incorporating the risk effect and the camp effect into the same regression equation, and tests the additive or substitution relationship between the two through interaction terms, thereby filling the gap between the two research paths in the existing literature. Second, using the net FDI inflow from the World Bank WDI as the dependent variable, all data are obtained directly from official sources to conduct genuine panel regression (coefficient, standard error, p-value, R-squared), rather than remaining at the level of descriptive statistics. Third, based on panel data of 22 major economies from 2006 to 2024, this paper employs the two-way fixed effects method to identify the differential impact of geopolitics on capital allocation, and multiple robustness tests—such as excluding the COVID period, pre-war sub-samples, and excluding G7—ensure that the conclusions are reliable.

2. Literature Review and Theoretical Assumptions

2.1 Measurement Evolution of Geopolitical Risks

The quantification of geopolitical risk has undergone an evolution from qualitative judgments to quantitative indices. In earlier research, Baker et al. tried to use the economic policy uncertainty (EPU) index to indirectly capture geopolitical factors, but the EPU is essentially a measurement of policy-level uncertainty, not a direct response to geopolitical risks, and there is a bias in reflecting purely geographic events such as military conflicts and territorial disputes [4].

This dilemma was not resolved until Caldara and Iacoviello constructed the GPR index [2]. By automating text mining of the electronic archives of 10 major English-language newspapers, the method retrieves the frequency of articles containing geopolitical tension keywords, constructs a monthly GPR index from 1900 to the present, and innovatively decomposes it into a threat component (GPRT) and an actual action component (GPRA). The threat component captures market expectations of potential conflict escalation, while the action component reflects actual military events that occurred. The GPRT value of 199.27 in 2022, well ahead of the GPRA's 124.80, suggests that market expectations of escalating conflicts are the main driver - this decomposition provides an important tool for analyzing the expected transmission mechanisms of capital flows. The advantages of text mining methods are: wide coverage, dating back to 1900, high frequency (monthly), and cross-national comparability through standardized processing.

Along with this, Bailey et al. based on the voting records of the United Nations General Assembly from 1946 to the present, using the Item Response Theory (IRT) model to estimate the ideal point of each country; the higher the ideal point, indicating that the political position of the country is to the US-led liberal order [3]. The model compresses tens of thousands of votes into a cross-country, year-round comparable political axis, removing the noise and issue specificity of a single vote. The ideal points in this sample range from -1.354 (China) to 2.696 (United States), covering the complete spectrum from opposition to the US-led liberal order to high alignment with that order, providing an operational tool for quantifying the camp effect. From indirect proxies to text mining, and from qualitative judgment to IRT model estimation, the root cause of this shift lies in the growth of computing power and the availability of large-scale digitized text archives. These advances enable researchers to process massive amounts of unstructured data and extract systematic geopolitical signals.

2.2 Total and Structural Effects

The aggregate effect refers to the phenomenon of capital flowing from high-risk countries to safe assets for hedging motives when geopolitical risks rise, and its core transmission mechanism is the change in risk premium. Rey's theory of the global financial cycle provides a theoretical basis for this effect: changes in global risk premiums drive the flow of capital between center countries and peripheral countries; GPR rises trigger flight to safety, and capital flows back from high-risk emerging markets to safe assets such as US Treasuries [5]. Aiyar et al. further found that geopolitical tensions led to the transfer of capital from emerging markets to developed markets, and this effect was particularly significant during periods of increased risk [1]. Empirical studies by Agoraki et al. confirmed a significant negative relationship between GPR and cross-

© 2026 The Author(s). Authors retain copyright in their articles.

border capital flows [6], while Feng et al. found that GPR also inhibited cross-border portfolio investment [7]. Together, these studies point to a conclusion: geopolitical risk systematically reduces the cross-border allocation of capital by raising the uncertainty premium.

The structural effect refers to the phenomenon that the distance between political positions leads to the reallocation of capital between different camps. The core transmission mechanism is the influence of ideological affinity on investment decisions. Aiyar et al. proposed a geo-economic fragmentation framework and found that the greater the distance between political positions, the less bilateral capital flows, and the fragmentation pressure is reshaping the global financial integration pattern [8]. Kempf et al. confirmed from the perspective of international lending that ideological distance significantly inhibits cross-border lending - banks are more inclined to lend to borrowers from countries with similar political positions [9]. Ciminelli et al. used the data of 30,000 fund holdings to provide micro evidence: when the GPR rose, the fund leaned toward the pro-China economy and stayed away from the pro-American economy, indicating that there was a clear camp preference for capital allocation [10]. Yang et al. further showed that GPR has a sustained negative impact on cross-border capital flows [11]. At present, the mainstream view is that the aggregate effect and the structural effect exist simultaneously. However, whether the two are additive or substitutive remains empirically untested within a unified framework integrating both mechanisms. This is precisely the gap this paper attempts to fill.

2.3 Theoretical Assumptions

Based on the above analysis, the existing research has the following limitations. First, the risk-hedging literature and the camp-allocation literature each tell only half the story, and no empirical test has incorporated the two mechanisms into a unified framework. Second, it remains impossible to answer the key question of whether the two effects are additive or substitutive. Some studies remain at the level of descriptive statistics due to the lack of variable data. They cannot complete rigorous regression analysis, making it difficult to provide credible coefficient estimates and statistical inferences. Based on these gaps, this paper explores the differential effects of GPR and political position distance on FDI using two-way fixed effects panel regression, and proposes the following three research hypotheses.

H1 (Risk - Hedge): Rising GPR results in reduced net inflows of FDI. It is based on the theory that rising risk premiums drive capital to seek safe havens.

H2 (Faction - Configuration): The higher the ideal point (i.e., closer to the US-dominated Western liberal order), the greater the net inflow of FDI. It is based on the theory that political distance inhibits capital flows.

H3 (Interaction Effect): The interaction term between GPR and the ideal point reflects the additive or substitution relationship between the two mechanisms. If the interaction term is negative, it supports the double-penalty hypothesis (additive effect). If the interaction term is positive, it suggests a substitution relationship, meaning that a pro-Western position has a hedging buffer effect in high-risk periods.

3. Study Design

3.1 Sample and Data Sources

This paper builds a panel dataset covering 22 major economies, 2006-2024. The sample was selected according to three principles: G7 developed economies (United States, Germany, United Kingdom, Japan, France, Italy, Canada, as the main source of capital); BRICS and G20 core emerging economies (China, Russia, India, Brazil, South Korea, Australia, Saudi Arabia, Turkey, Mexico, Indonesia, Argentina, South Africa, Egypt, as important capital destinations); Countries involved in major geopolitical conflicts (Ukraine, Israel). Covering more than 80% of global GDP, the sample is broadly representative. All data are obtained directly from official sources without any interpolation or manual values. FDI, GDP and financial deepening data are from the World Bank's WDI database [12], the GPR index is from the official websites of Caldara and Iacoviello [2], and the ideal point is estimated from Harvard Dataverse by Bailey et al. [3].

© 2026 The Author(s). Authors retain copyright in their articles.

3.2 Variable Definitions

The variable system includes dependent variables (net FDI inflows as a proportion of GDP), core independent variables (GPR, ideal points), interaction terms, and control variables (ln (GDP), financial deepening). The specific definitions are shown in Table 1. As can be seen in Table 1, the dependent variable FDI comes from the World Bank's WDI indicator BX.KLT.DINV.WD.GD.ZS, which measures the scale of a country's attraction of foreign direct investment. A positive value indicates net capital inflows, while a negative value indicates net capital outflows [12]. The core independent variable GPR is the annual average of the national monthly indexes constructed by Caldara and Iacoviello, and the higher the value, the greater the geopolitical risk facing the country [2]. The ideal point is the UN voting ideological coordinate estimated by Bailey et al. The higher the value, the closer the political position is to the US-led liberal order [3]. Among the control variables, ln(GDP) controls for the economic scale effect—countries with larger economies usually attract more capital. Fin (domestic credit as a proportion of GDP) controls for the depth of financial development—the more developed the financial system, the higher the efficiency of capital allocation. It should be noted that the institutional quality variables were removed from the model because the World Bank WGI database API is no longer available for the relevant indicators.

Table 1: Variable definitions and data sources

Variable Category	Variable Name	Symbol	Definition	Data Source
Explained Variable	FDI Net Inflows	FDI	FDI net inflows (% of GDP)	World Bank WDI [12]
Explanatory variable	Geopolitical Risk	GPR	Country-level GPR monthly index annual mean	Caldara & Iacoviello [2]
	Ideal Point	IdealPoint	UN voting ideological coordinate	Bailey et al. [3]
Control Variable	Economic Scale	ln(GDP)	Natural logarithm of GDP	World Bank WDI [12]
	Financial Depth	Fin	Domestic credit (% of GDP)	World Bank WDI [12]
Interaction	Risk x Position Interaction	GPR x IdealPoint	Mean-centered GPR x IdealPoint product	Constructed by this study

3.3 Empirical Model Setting

In this paper, two regression models are set up. Model (1) is the benchmark model, which aims to test the independent influence of GPR and ideal point on FDI, that is, whether GPR is significantly negative (validation H1) and ideal point is significantly positive (validation H2) after controlling for each other's effects and other control variables. The inclusion of two core independent variables at the same time is to avoid the bias of missing variables - if only GPR is included without the ideal point, the GPR coefficient may absorb the influence of political position distance. Model (2) is the interaction effect model, and the interaction term between GPR and ideal point is added on the basis of the benchmark model to test whether the two mechanisms are additive or substitutive (verify H3). GPR and IdealPoint are mean-centered before constructing the interaction term to mitigate multicollinearity. Both models control for the fixed effects of the country (absorbing the national characteristics that do not change with time, such as geographical location, resource endowment) and the fixed effects of the year (absorbing the time shocks that do not change with the country, such as the global financial crisis and the epidemic), and standard errors are clustered at the country level.

$$FDI = \alpha + \beta_1 \cdot GPR + \beta_2 \cdot IdealPoint + \gamma_1 \cdot ln(GDP) + \gamma_2 \cdot Fin + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

Among them, mu_i is the country fixed effect, lambda_t is the year fixed effect, epsilon_it is the random error term, and standard errors are clustered at the country level. beta_1 < 0 confirms H1, beta_2 > 0 confirms H2.

$$FDI = \alpha + \beta_1 \cdot GPR + \beta_2 \cdot IdealPoint + \beta_3 \cdot (GPR \times IdealPoint) + \gamma \cdot Controls + \mu_i + \lambda_t + \varepsilon_{it} \quad (2)$$

beta_3 is the core coefficient of H3. If beta_3 < 0, the double penalty hypothesis (additive effect) is supported; if beta_3 > 0, a substitution relationship is suggested (a pro-Western position has a hedging buffer effect in high-risk periods).

© 2026 The Author(s). Authors retain copyright in their articles.

3.4 Descriptive Statistics

Table 2: Descriptive statistics

Variable	N	Mean	Std. Dev.	Min	Max
FDI (%GDP)	383	2.205	1.593	-0.894	12.000
GPR	383	0.404	0.562	0.014	3.609
IdealPoint	383	0.583	0.934	-1.354	2.696
ln(GDP)	383	28.077	1.132	25.234	31.009
Fin (%)	383	91.868	51.916	11.910	219.126

Table 2: a total of 383 observations in the sample, covering data from 22 countries from 2006 to 2024. The average FDI was 2.205%, with a standard deviation of 1.593, ranging from -0.894% (net capital outflows) to 12.000%, with significant cross-country differences, reflecting the huge differentiation of countries' ability to attract foreign investment. The GPR averaged 0.404, but reached a maximum of 3.609 (Russia 2022), nearly nine times the average, reflecting the high concentration and time-varying nature of geopolitical risks. The ideal point average is 0.583, ranging from -1.354 to 2.696, with a standard deviation of 0.934, covering the complete spectrum from opposition to the US-dominated Western liberal order to a high approximation to that order, providing sufficient cross-sectional variation for testing the camp effect. The control variables of ln (GDP) and Fin also exhibit reasonable variability.

4. Empirical Results and Analysis

Table 3: Benchmark regression results

Variable	Model (1) Baseline	Model (2) Interaction
GPR	-1.031***	-1.546***
	(0.248)	(0.226)
IdealPoint	-0.042	0.079
	(0.419)	(0.483)
GPR x IdealPoint	—	0.776***
		(0.234)
ln(GDP)	-0.134	-0.083
	(0.550)	(0.499)
Fin	-0.006	-0.004
	(0.006)	(0.006)
Country FE / Year FE	Yes / Yes	Yes / Yes
R2	0.501	0.509
N	383	383

*Note: clustered standard errors in brackets. * * * p < 0.01, * * p < 0.05, * p < 0.1.*

In Table 3, the GPR coefficient in model (1) is -1.031 ($p < 0.01$), that is, for every unit of increase in the GPR, the net FDI inflow decreases by 1.03 percentage points, confirming H1. Taking Russia in 2022 as an example, its GPR soared from 0.82 to 3.69, an increase of about 2.87 units, corresponding to a decrease in FDI of about 2.95 percentage points - the actual FDI plummeted from a positive value to a negative value, which is highly consistent with the forecast. Among them, the GPR data were taken from the monthly index at the country level published by the official websites of Caldara and Iacoviello (2022) [2], and the FDI data were taken from the World Bank WDI database (indicator code BX.KLT.DINV.WD.GD.ZS). It should be noted that the above Russian case reflects only the independent effect of GPR and is not related to the ideal point coefficient. The independent effect of the ideal point itself is separately characterized by its coefficient: the ideal point coefficient in model (1) is -0.042 ($p = 0.920$), which is not significant, indicating that the political position distance alone does not affect FDI without considering geopolitical risks, and H2 is not supported as an independent effect. One possible explanation is that in normal times (low GPR), economic fundamentals (market size, level of financial development) dominate capital flow decisions, and differences in the political stances of investment destinations are masked by economic motivations. Only when geopolitical risk rises does position distance amplify the impact on FDI through the risk perception channel - and whether this adjustment mechanism holds is the question to be tested by the interaction term in Model (2).

After the interaction term was added to Model (2), the GPR coefficient became -1.546 ($p < 0.01$), which was stronger than the benchmark model, because the interaction term absorbed part of the variation caused by the heterogeneity of the political stance in the GPR effect. The interaction term coefficient was 0.776 ($p <$

© 2026 The Author(s). Authors retain copyright in their articles.

0.01), which was significantly positive, contrary to the double-penalty hypothesis expected under H3. This result shows that countries with pro-Western political stances experience smaller FDI shocks when GPR rises, and the two mechanisms exhibit a substitution relationship rather than an additive relationship. Specifically, when GPR rises by 1 unit, the negative effect on FDI is reduced by 0.776 percentage points for each unit increase in the ideal point (i.e., more pro-Western). In other words, the pro-Western stance has hedging properties in times of high geopolitical risk, partially cushioning the risk shock. Neither the control variables ln (GDP) nor Fin are significant, indicating that cross-country differences in economic size and financial depth have been absorbed by national characteristics after controlling for national fixed effects. R-squared increased from 0.501 to 0.509, indicating that the interaction term provides additional explanatory power.

5. Robustness Testing

Table 4: Robustness test

Test	GPR Coef.	GPR p-value	Interaction Coef.	Interaction p-value	N
R1: Excluding COVID (2020-2021)	-1.623***	0.000	0.836***	0.002	343
R2: Pre-war (2006-2021)	-1.439**	0.045	0.748*	0.052	327
R3: Excluding G7	-1.900***	0.000	0.935***	0.009	267
Baseline (Reference)	-1.546***	0.000	0.776***	0.001	383

*Note: All models contain country fixed effects, year fixed effects, and clustered standard errors. * * * p < 0.01, * * p < 0.05, * p < 0.1.*

As shown in Table 4, in order to test the robustness of the conclusions, this paper carries out three tests. After R1 excluded the COVID period (2020-2021), the GPR effect was stronger (-1.623, $p < 0.01$), and the interaction term was more significant (0.836, $p < 0.01$), indicating that the abnormal capital flow policies during the pandemic constituted noise, and the true effect of geopolitics becomes clearer after excluding this period. GPR is still significant (-1.439, $p = 0.045$) in the R2 pre-war sub-sample (2006-2021), and the interaction term is significant at the 10% level (0.748, $p = 0.052$), indicating that the impact of geopolitics on capital allocation is not only driven by the 2022 war, but a persistent structural phenomenon - geopolitical shifts began well before the Russo-Ukrainian War. After excluding the G7 advanced economies in R3, the GPR effect is further strengthened (-1.900, $p < 0.01$), and the interaction term is more significant (0.935, $p < 0.01$), which confirms that emerging markets face higher risk premiums, and the buffering effect of political positions is more obvious in emerging markets - these countries are more vulnerable in international capital markets, and the marginal impact of political positions is greater.

6. Conclusion

Based on the panel data of 22 major economies from 2006 to 2024, this paper uses two-way fixed-effects regression to examine the differential impact of geopolitical risks and political position distances on FDI. The study found that: First, GPR has a strong negative effect on FDI. For every unit increase in GPR, the net inflow of FDI decreases by 1.55 percentage points ($p < 0.01$). After excluding COVID and focusing on emerging markets, the effect is stronger, confirming the risk-hedging preference hypothesis. Second, the independent effect of the ideal point is not significant, but the interaction between the GPR and the ideal point is significantly positive (coefficient 0.78, $p < 0.01$), indicating that pro-Western countries experience smaller FDI shocks under high geopolitical risk, and the two mechanisms exhibit a substitution relationship rather than an additive relationship - the pro-Western stance has a safe-haven property in high-risk periods, partially buffering the risk shock. This finding corrects the simple double-penalty hypothesis in the literature and reveals more complex mechanisms through which geopolitics influences capital allocation.

This study has the following limitations: First, the dependent variable is national FDI rather than bilateral capital flows, which cannot capture changes in capital flows between specific country pairs. In the future, the IMF CPIS bilateral data can be used to deepen the analysis. Second, institutional quality variables were not included due to data availability constraints, and alternative indicators may be found in future research. Third, the measure of distance between political positions relies on UN voting records and may not fully reflect informal diplomatic relations. Based on the above findings, it is recommended that emerging market countries incorporate GPR into the macroprudential framework and establish a geopolitical stress testing mechanism; When balancing strategic autonomy and capital friendliness, China should be wary of the cumulative impact

© 2026 The Author(s). Authors retain copyright in their articles.

that excessive political deviation may have on capital inflows and global governance participation. Geopolitical considerations should be incorporated when assessing the prospects of attracting foreign investment. Future research can further deepen the analysis using micro enterprise investment data and bilateral political relationship indicators constructed by NLP.

References

[1] Aiyar, S., Chen, J., Ebeke, C., et al. (2023b). Geopolitics and Capital Flows. IMF WEO, Ch. 4.

[2] Caldara, D., & Iacoviello, M. (2022). Measuring Geopolitical Risk. AER, 112(4), 1194-1225.

[3] Bailey, M. A., Strezhnev, A., & Voeten, E. (2017). Estimating Dynamic State Preferences from UN Voting Data. J. Conflict Resolution, 61(2), 430-456.

[4] Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring Economic Policy Uncertainty. QJE, 131(4), 1593-1636.

[5] Rey, H. (2015). Dilemma not Trilemma: The Global Financial Cycle. NBER WP 21162.

[6] Agoraki, M. E. K., Kolympiris, C., & Papathanasiou, V. (2024). Geopolitical Risk and Capital Flows. J. Int. Financial Markets, Institutions & Money, 93, 101983.

[7] Feng, Y., Gong, X., & Yang, J. (2023). GPR and Cross-Border Portfolio Investment. J. Int. Money & Finance, 130, 102768.

[8] Aiyar, S., Presbitero, A., & Ruta, M. (2023a). Geoeconomic Fragmentation and the Future of Multilateralism. IMF WP/23/190.

[9] Kempf, E., Margotta, P., & Tsoutsoura, M. (2023). Ideological Distance and International Lending. WP.

[10] Ciminelli, G., Raabe, D., & Zhao, X. (2026). GPR and Mutual Fund Portfolio Reallocation. IMF WP.

[11] Yang, J., Feng, Y., & Chen, S. (2026). GPR and Cross-Border Capital Flows. J. Int. Economics.

[12] World Bank (2026). World Development Indicators. Indicator: BX.KLT.DINV.WD.GD.ZS.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.118128
Published by: Zeus Press

VAT Credit Refunds and Firms' Export Development Evidence from Chinese A-Share Listed Companies

Ken Peng*

University of Shanghai for Science and Technology, Shanghai, China

**Corresponding author: Ken Peng.*

Abstract

This study examines whether China's value-added tax (VAT) credit refund policy supports firms' export development through financing relief and productivity improvements. Using an unbalanced panel of 6014 firm-year observations covering 821 A-share listed firms during 2013–2023, we estimate difference-in-differences models with firm and year fixed effects. The fully controlled policy coefficient on log overseas operating revenue is 0.2202 and is significant at the 1% level. Event-study estimates do not reject parallel pre-policy trends, while placebo, propensity score matching, weighting, and sample-period checks support the positive association. Additional estimates show an improvement in firm resilience. Mechanism regressions yield positive indirect effects through reduced financing constraints and higher total factor productivity, with firm-clustered bootstrap confidence intervals excluding zero. The point estimate is larger for non-state-owned firms, although the difference between ownership groups is not formally tested. The findings are consistent with liquidity and productivity channels supporting overseas business activity. Because overseas revenue is a scale-based proxy, the results do not directly establish improvements in export product quality or domestic value added.

Keywords

vat credit refunds, export development, financing constraints, total factor productivity, difference-in-differences

1. Introduction

High-quality export development requires firms to sustain foreign-market participation while improving productive capabilities and resilience. Export revenue is an important part of this process, but it does not capture all dimensions of competitiveness. Ceglowski [1] shows that assessments based on gross exports can differ from those based on domestic value added. This distinction is relevant when firm-level financial data are used to study trade-related outcomes: stronger overseas sales may reflect greater scale, changing prices, or a different composition of activities rather than product upgrading alone.

Financial frictions provide one explanation for uneven international performance across firms. Entering and serving foreign markets requires expenditure on production, distribution, customer acquisition, and working capital before sales proceeds are received. Bellone et al. [2] find that financially healthier firms are more likely to become exporters, whereas Greenaway et al. [3] emphasize that export participation can itself

© 2026 The Author(s). Authors retain copyright in their articles.

improve financial health. These findings suggest that liquidity and exporting may be jointly determined, making policy-related changes in available funds a useful setting for examining their relationship.

VAT credit refunds release funds tied up in unutilized input tax credits. Unlike a reduction in the statutory tax rate, a refund changes the timing and availability of liquidity associated with tax credits. It may allow firms to maintain production, reduce reliance on costly external finance, and undertake productivity-enhancing expenditure. Related evidence from China links VAT export rebates to product upgrading through credit constraints [4]. However, export rebates and refunds of accumulated input VAT credits are distinct policy instruments and should not be treated as interchangeable. Research on China's VAT credit refund reform also documents a reduction in corporate financing constraints [5].

This study examines VAT credit refunds and firms' export development using 6014 firm-year observations from 821 Shanghai and Shenzhen A-share listed firms over 2013–2023. The analysis combines firm and year fixed effects, an event study, placebo exercises, matching and weighting checks, mechanism regressions, and ownership-specific estimates. The contribution is to connect overseas operating revenue with both financing constraints and total factor productivity in a common empirical framework, while using firm resilience as a complementary outcome. The evidence concerns overseas business performance; it does not provide a direct measure of product quality, export sophistication, or supply chain resilience.

The baseline policy coefficient is positive and remains statistically significant across the reported alternative specifications. The mechanism estimates are consistent with financing relief and productivity improvements, and the ownership results suggest a larger point estimate among non-state-owned firms. These findings are interpreted subject to the policy assignment, parallel trends, and mediation assumptions discussed below. Section 2 develops the institutional background and hypotheses, Section 3 describes the research design, Section 4 presents the results, and Section 5 concludes.

2. Policy Background and Research Hypotheses

The proposed transmission mechanism links the release of accumulated tax credits to firms' financing conditions and productive capacity. The policy background also matters for interpreting comparisons over a period in which refund eligibility expanded.

2.1 Institutional Background

China's 2018 reform permitted refunds of end-of-period VAT credits for selected advanced manufacturing industries, modern services such as research and development, and eligible power-grid enterprises. The national notice specified industry coverage and additional eligibility conditions, including tax credit ratings [6]. Industry membership therefore represents potential exposure rather than proof that every covered firm actually received a refund. This distinction motivates interpreting an industry-based DID indicator as policy exposure rather than the amount of refund received.

The policy subsequently broadened. An incremental VAT credit refund system was introduced across industries from April 2019, subject to eligibility requirements [7]. In 2022, refund support expanded further for qualifying small and micro enterprises and specified industries [8]. Consequently, a long post-reform window may reflect evolving relative exposure rather than a permanent comparison between recipients and firms wholly unaffected by refund policies. Firm and year fixed effects absorb persistent firm differences and common annual shocks, but they do not automatically remove differential exposure to later reforms.

2.2 Liquidity and Financing Constraints

Unrefunded input VAT credits tie up cash that firms could otherwise use for working capital. Returning these credits can reduce the need for short-term borrowing and ease the funding of production and overseas orders. For firms facing a wedge between internal and external financing costs, the release of liquidity may support activities that would otherwise be delayed or scaled back. This reasoning connects the financial barriers to exporting documented in the trade literature with the financing relief associated with VAT credit refunds.

H1. Exposure to the VAT credit refund policy is associated with higher overseas operating revenue.

© 2026 The Author(s). Authors retain copyright in their articles.

If financing relief is part of the transmission process, policy exposure should reduce the WW financing constraints index, while tighter constraints should be associated with weaker overseas performance conditional on policy exposure and firm characteristics. The analysis treats this channel as financing-constraint relief; the WW index is not a direct measure of working capital turnover or operating-cycle efficiency.

H2. Reduced financing constraints mediate the positive relationship between VAT credit refund exposure and overseas operating revenue.

2.3 Productivity and Ownership

Liquidity may also affect longer-term productive capacity. Firms able to fund equipment maintenance, process improvement, and complementary inputs can improve production efficiency and their ability to meet foreign demand. VAT-related trade research links additional cash flow to upgrading investment, suggesting a possible productivity channel. Whether these opportunities translate into higher productivity is an empirical question, since refunds may also be used to repay liabilities or accumulate cash.

H3. Higher total factor productivity mediates the positive relationship between VAT credit refund exposure and overseas operating revenue.

The marginal value of released liquidity may differ by ownership. Where non-state-owned firms face tighter external financing constraints, a refund could have a larger influence on their overseas activity. Ownership-specific regressions therefore provide an exploratory extension. A significant coefficient in one group and an insignificant coefficient in another, however, do not by themselves establish a statistically significant difference between groups.

3. Research Design

The empirical analysis links firm-year financial information, overseas operating revenue, and industry-based policy exposure. The model specifications distinguish the overall policy coefficient from the statistical evidence on potential channels.

3.1 Sample and Data

The sample comprises Shanghai and Shenzhen A-share listed firms over 2013–2023. Firm financial information is matched with overseas operating revenue and policy-related industry classifications. Observations with missing core variables are excluded, and continuous variables are winsorized within each year at the 1st and 99th percentiles. The final unbalanced panel contains 6014 firm-year observations from 821 firms. Coverage differs across years, with relatively few observations in 2013 and 2014, so early event-study estimates warrant particular caution.

3.2 Variables

The dependent variable, Overseas_Rev, is the log overseas operating revenue measure in the estimation sample. It captures the scale of overseas business activity and serves as a limited indicator of export development. Overseas revenue may include activities beyond customs-recorded exports and does not identify product quality, domestic value added, or export prices separately. The analysis therefore interprets changes in this outcome as changes in overseas operating performance.

The core explanatory variable is DID, the interaction between Treat and Post. Treat equals one when a firm's industry in the current year falls within the designated policy coverage and zero otherwise. Post indicates the post-implementation period. Since current-year industry membership is used, changes in industry classification may alter exposure within a firm. This coding differs from a treatment group fixed using pre-policy industry membership.

WW measures financing constraints, with larger values indicating tighter constraints, following the interpretation of the Whited–Wu index [9]. TFP_LP measures total factor productivity using the Levinsohn–Petrin approach, which uses intermediate inputs to address the correlation between input choices and unobserved productivity [10]. Resilience1 is an alternative outcome measuring firm resilience, not supply

© 2026 The Author(s). Authors retain copyright in their articles.

chain resilience. The control variables are tax refund intensity (Rebates), firm size (Size), the debt-to-asset ratio (Lev), and return on assets (ROA). Table 1 summarizes the variables.

Table 1: Variable definitions

Variable type	Variable name	Symbol	Definition and description
Dependent variable	Overseas operating performance	Overseas_Rev	Log of overseas operating revenue
Core explanatory variable	Policy shock	DID	Treat×Post; Treat is determined by whether the firm's industry in the current year is covered by the policy
Mechanism variable	Financing constraints	WW	WW financing constraints index
Mechanism variable	Total factor productivity	TFP_LP	Total factor productivity estimated using the LP method
Alternative dependent variable	Firm resilience	Resilience1	Firm resilience measure
Controls	Tax refund intensity	Rebates	Firm tax refund intensity measure
Controls	Firm size	Size	Firm size measure
Controls	Debt-to-asset ratio	Lev	Firm leverage
Controls	Return on assets	ROA	Firm profitability

Note: Variable symbols are retained throughout. Overseas_Rev is a scale-based overseas revenue measure. Resilience1 denotes firm resilience.

3.3 Baseline Specification

The baseline difference-in-differences model includes firm and year fixed effects:

$$\text{Overseas_Rev}_{it} = \alpha_0 + \beta \text{DID}_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

Here, i indexes firms and t indexes years. X contains Rebates, Size, Lev, and ROA; μ and λ denote firm and year fixed effects, respectively. The coefficient β measures the differential change in log overseas operating revenue associated with policy exposure, conditional on the included controls. Standard errors are clustered at the firm level to allow serial dependence within firms, an important concern in panel DID inference [11]. Because policy exposure is defined by industry, residual dependence across firms within industries remains an additional inference concern.

A causal interpretation requires parallel counterfactual trends, no relevant anticipation, and the absence of unaccounted differential shocks or spillovers. Industry-specific demand movements and subsequent refund reforms could challenge these conditions. Moreover, contemporaneous controls such as Rebates may themselves respond to the policy. The specification without firm controls and the fully controlled specification are therefore both reported, rather than assuming that their coefficients necessarily identify the same total effect.

3.4 Mechanism and Ownership Specifications

For each mechanism variable M, the analysis estimates a policy-to-mediator equation and an outcome equation conditional on the mediator:

$$M_{it} = a_0 + a_1 \text{DID}_{it} + \delta' X_{it} + \mu_i + \lambda_t + \nu_{it} \quad (2)$$

$$\text{Overseas_Rev}_{it} = c_0 + c' \text{DID}_{it} + bM_{it} + \theta' X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (3)$$

M denotes WW or TFP_LP in separate models. The estimated indirect effect is $a_1 b$. Percentile and bias-corrected confidence intervals are obtained from 1000 bootstrap replications clustered by firm. An interval excluding zero supports a nonzero indirect association under the specified linear model. It does not by itself establish causal mediation: interpreting the product as a causal pathway requires additional assumptions about mediator–outcome confounding [12].

Ownership heterogeneity is examined by estimating the baseline equation separately for state-owned and non-state-owned enterprises:

$$\text{Overseas_Rev}_{it} = \alpha_0 + \beta^g \text{DID}_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}, g \in \{\text{SOE}, \text{non} - \text{SOE}\} \quad (4)$$

© 2026 The Author(s). Authors retain copyright in their articles.

The group-specific coefficients describe the estimated policy relationship within each ownership category. Their point estimates and within-group significance are reported without treating them as a formal test of cross-group equality.

4. Empirical Results

The results are presented in the order of descriptive evidence, baseline estimates, identification diagnostics, robustness checks, potential mechanisms, and ownership differences. The discussion distinguishes the evidence on overseas revenue from broader claims about product quality and export upgrading.

4.1 Descriptive Statistics and Diagnostic Checks

Table 2 reports substantial variation in overseas business activity: Overseas_Rev has a mean of 19.376 and a standard deviation of 2.038. The mean of DID is 0.599, indicating that 59.9% of firm-year observations are coded as policy-exposed. The means of WW and TFP_LP are −1.017 and 8.650, respectively. Resilience1 has a mean of 0.011 and a standard deviation of 0.102. These distributions provide variation for the subsequent regressions, but do not establish a policy effect.

Table 2: Descriptive statistics

Variable	Observations	Mean	Std. dev.	Min.	Max.	P1	P99
Overseas_Rev	6014	19.376	2.038	12.164	24.094	13.609	23.630
DID	6014	0.599	0.490	0.000	1.000	0.000	1.000
WW	6014	-1.017	0.061	-1.207	-0.863	-1.179	-0.885
TFP_LP	6014	8.650	0.892	6.521	12.220	6.821	11.202
Rebates	6014	0.011	0.013	0.000	0.078	0.000	0.064
Size	6014	22.213	1.046	20.065	26.215	20.186	25.261
Lev	6014	0.408	0.180	0.061	0.953	0.082	0.861
ROA	6014	0.029	0.080	-0.528	0.235	-0.325	0.183
Resilience1	6014	0.011	0.102	-0.483	0.492	-0.286	0.346

Note: P1 and P99 denote the 1st and 99th percentiles, respectively.

The correlation between DID and Overseas_Rev is 0.050 and is significant at the 1% level (Table 3). This positive unconditional association is consistent with the proposed relationship but may reflect other firm characteristics. The baseline variance inflation factors are all below 1.5, with a mean of 1.196 (Table 4). These values do not indicate serious multicollinearity in the baseline covariate set. They do not cover the additional mechanism variables: WW, TFP_LP, and Size exhibit relatively high pairwise correlations, which should be considered when interpreting the mechanism regressions.

Table 3: Correlation matrix

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Overseas_Rev	1								
(2) DID	.050***	1							
(3) WW	-.459***	-.002	1						
(4) TFP_LP	.540***	.059***	-.711***	1					
(5) Rebates	.349***	.042***	.005	.102***	1				
(6) Size	.514***	.071***	-.813***	.817***	-.041***	1			
(7) Lev	.229***	.038***	-.161***	.389***	.010	.402***	1		
(8) ROA	.102***	-.070***	-.379***	.164***	.067***	.071***	-.347***	1	
(9) Resilience1	.012	.070***	-.083***	.051***	.007	-.010	-.081***	.244***	1

*Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.*

Table 4: Variance inflation factors

Variable	VIF	1/VIF	Variable	VIF	1/VIF
Lev	1.441	0.694	Size	1.283	0.779
ROA	1.227	0.815	DID	1.014	0.986
Rebates	1.013	0.987	Mean	1.196	—

Note: VIFs apply to DID and the control variables in the baseline model.

© 2026 The Author(s). Authors retain copyright in their articles.

4.2 Baseline Estimates

Table 5 presents the baseline results. With firm and year fixed effects only, the DID coefficient is 0.3921 with a standard error of 0.0683. Adding the firm controls reduces the coefficient to 0.2202, with a standard error of 0.0592; both estimates are significant at the 1% level. The positive coefficient supports H1 as an empirical association between policy exposure and overseas operating revenue. The fully controlled estimate is expressed in log points, rather than being read directly as a percentage change.

Table 5: VAT credit refund exposure and overseas operating revenue

Variable	(1)	(2)
DID	0.3921***	0.2202***
	(0.0683)	(0.0592)
Rebates		22.6524***
		(2.0445)
Size		0.8549***
		(0.0546)
Lev		0.1693
		(0.1793)
ROA		0.7739***
		(0.2070)
Constant	19.1415***	-0.0878
	(0.0409)	(1.1908)
Firm fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	6014	6014
Adjusted R^2	0.8746	0.8966

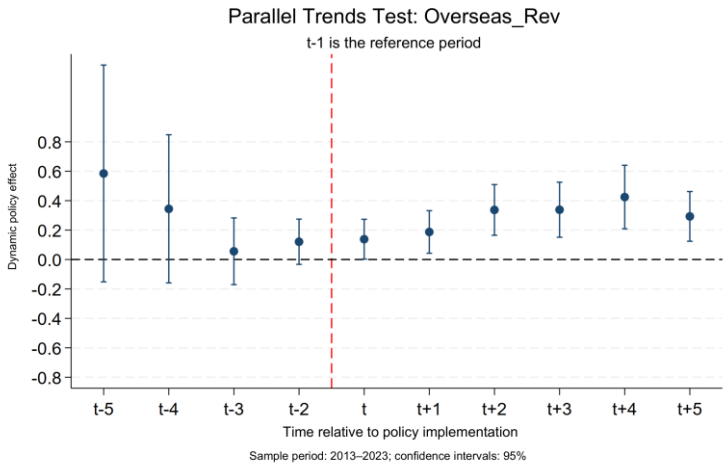
*Note: The dependent variable is Overseas_Rev. Firm-clustered robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

Rebates, Size, and ROA have positive coefficients significant at the 1% level, whereas Lev is not significant at conventional levels. Adjusted R^2 increases from 0.8746 to 0.8966 when the controls are included. This improvement in fit does not resolve identification concerns, but the positive DID coefficient in both specifications shows that the result is not confined to the specification without firm controls.

4.3 Event-study and Placebo Evidence

Figure 1 normalizes the period immediately before policy implementation to zero. The pre-policy confidence intervals include zero, and the joint pre-policy test has a p-value of 0.2519. Thus, the test does not reject equal pre-policy trends. This finding is compatible with the identifying assumption, although wide early confidence intervals and sparse initial-year coverage limit the power to detect departures. Post-policy point estimates are positive and generally remain above zero, with intervals generally excluding zero from the first subsequent period onward.

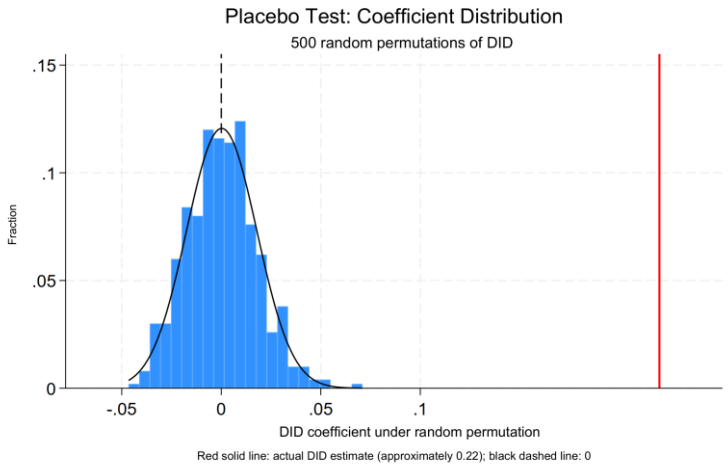
Figure 1: Event-study estimates of overseas operating revenue



© 2026 The Author(s). Authors retain copyright in their articles.

Figure 2 summarizes 500 random permutations of the policy indicator. The placebo coefficients are concentrated around zero, while the actual DID coefficient of approximately 0.22 lies to the right of the placebo distribution. This pattern reduces concern that an arbitrary assignment mechanically reproduces the baseline result. The graphical evidence does not supply an exact empirical p-value and cannot rule out systematic industry-specific confounding.

Figure 2: Distribution of placebo policy coefficients



4.4 Matching Weighting and Alternative Outcomes

Propensity score methods address differences in observed characteristics between treatment and comparison units, subject to overlap and an appropriate assignment model [13]. Table 6 reports the matching and weighting estimates. With controls included, the matched-sample DID coefficient is 0.2280 and the weighted coefficient is 0.2210; both are significant at the 1% level and close to the baseline estimate of 0.2202. The corresponding estimation samples contain 4868 observations. Matching and weighting do not address selection on unobserved characteristics or exposure to other policies.

Table 6: Matching and weighting estimates

Variable	Before matching	After matching	Matched + controls	Weighted	Weighted + controls
DID	0.2200***	0.4180***	0.2280***	0.3400***	0.2210***
	(3.716)	(5.807)	(3.726)	(4.759)	(3.539)
Rebates	22.6520***		24.1840***		22.6650***
	(11.080)		(10.723)		(9.639)
Size	0.8550***		0.8610***		0.7740***
	(15.647)		(14.563)		(7.103)
Lev	0.1690		0.3850**		0.6280*
	(0.944)		(2.059)		(1.901)
ROA	0.7740***		0.7090***		0.4780
	(3.739)		(3.298)		(1.034)
Constant	-0.0880	19.2690***	-0.2630	19.3320***	1.6510
	(-0.074)	(460.059)	(-0.202)	(777.926)	(0.700)
Controls	Yes	No	Yes	No	Yes
Firm/year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	6014	4868	4868	4868	4868
R^2	0.897	0.873	0.897	0.893	0.912

*Note: Parentheses contain t-statistics, not standard errors. All models include firm and year fixed effects. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

Table 7 examines a complementary outcome and an alternative sample period. When Resilience1 is the dependent variable, the DID coefficient is 0.0208 with a standard error of 0.0066, significant at the 1% level. This result concerns firm resilience and does not establish supply chain resilience or validate overseas revenue

© 2026 The Author(s). Authors retain copyright in their articles.

as a direct product-quality measure. Excluding 2018 yields a DID coefficient of 0.2433 with a standard error of 0.0642, also significant at the 1% level. The reported reduced-sample result is consistent in sign with the baseline estimate.

Table 7: Firm resilience and alternative sample period

Variable	Firm resilience	Excluding 2018
DID	0.0208***	0.2433***
	(0.0066)	(0.0642)
Rebates	-0.2354	24.5651***
	(0.1952)	(2.4005)
Size	-0.0299***	0.8527***
	(0.0055)	(0.0572)
Lev	0.0315	0.3090*
	(0.0209)	(0.1851)
ROA	0.4643***	1.0186***
	(0.0403)	(0.2490)
Constant	0.6385***	-0.1328
	(0.1201)	(1.2483)
Firm/year fixed effects	Yes	Yes
Observations	6014	5269
R^2	0.1036	0.8955

*Note: Column (1) uses Resilience1; column (2) uses Overseas_Rev and excludes 2018. Firm-clustered robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

4.5 Financing Constraints and Productivity Channels

Table 8 reports the mechanism regressions. In column (1), the DID coefficient on WW is −0.0054 and is significant at the 5% level. In column (2), WW has a coefficient of −2.7272, significant at the 1% level, while the DID coefficient remains positive at 0.2054. The pattern is consistent with H2: policy exposure is associated with lower financing constraints, and lower constraints are associated with stronger overseas operating performance.

Table 8: Financing constraints and productivity regressions

Variable	WW	Overseas_Rev	TFP_LP	Overseas_Rev
DID	-0.0054**	0.2054***	0.0718**	0.1764***
	(0.0025)	(0.0588)	(0.0300)	(0.0564)
WW		-2.7272***		
		(0.5008)		
TFP_LP				0.6089***
				(0.0826)
Rebates	0.0323	22.7405***	4.4992***	19.9126***
	(0.0591)	(2.0412)	(0.9732)	(1.9829)
Size	-0.0577***	0.6974***	0.5255***	0.5349***
	(0.0016)	(0.0601)	(0.0238)	(0.0636)
Lev	0.0110*	0.1993	0.2927***	-0.0090
	(0.0063)	(0.1771)	(0.0860)	(0.1688)
ROA	-0.1644***	0.3255	1.1487***	0.0744
	(0.0098)	(0.2214)	(0.1248)	(0.2163)
Constant	0.2685***	0.6445	-3.2691***	1.9029*
	(0.0356)	(1.1869)	(0.5139)	(1.1364)
Firm/year fixed effects	Yes	Yes	Yes	Yes
Observations	6014	6014	6014	6014
R^2	0.8317	0.8977	0.9299	0.9015

*Note: Columns (1)–(4) use WW, Overseas_Rev, TFP_LP, and Overseas_Rev, respectively. Firm-clustered robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

For productivity, the DID coefficient on TFP_LP is 0.0718 and is significant at the 5% level. When TFP_LP enters the overseas-revenue equation, its coefficient is 0.6089 and is significant at the 1% level; the remaining DID coefficient is 0.1764 and is also significant at the 1% level. This pattern is consistent with H3 and a

© 2026 The Author(s). Authors retain copyright in their articles.

productivity channel, although the contemporaneous mediator regressions cannot exclude reverse causality or common shocks to productivity and overseas sales.

Table 9 reports the bootstrap results. The WW indirect effect is 0.0147152, with a percentile 95% confidence interval of [0.0007610, 0.0315808] and a bias-corrected interval of [0.0011601, 0.0317664]. The TFP_LP indirect effect is 0.0437362, with corresponding intervals of [0.0075793, 0.0836514] and [0.0112591, 0.0908695]. All four intervals exclude zero. Together with the remaining positive DID coefficients, these estimates are consistent with partial mediation within each separate model.

Table 9: Firm-clustered bootstrap mechanism estimates

Mechanism variable	Path	Observed coefficient	Bootstrap SE	Percentile 95% CI	Bias-corrected 95% CI	Conclusion
WW	a	-0.0053957	0.0026701	[-0.0109360, -0.0003004]	[-0.0109295, -0.0000276]	Excludes 0
WW	b	-2.7272214	0.5056615	[-3.7453980, -1.7412440]	[-3.7077290, -1.7352180]	Excludes 0
WW	ab	0.0147152	0.0078329	[0.0007610, 0.0315808]	[0.0011601, 0.0317664]	Excludes 0
TFP_LP	a	0.0718231	0.0297719	[0.0125791, 0.1286949]	[0.0168722, 0.1309467]	Excludes 0
TFP_LP	b	0.6089431	0.0820797	[0.4504982, 0.7795874]	[0.4486414, 0.7788323]	Excludes 0
TFP_LP	ab	0.0437362	0.0194064	[0.0075793, 0.0836514]	[0.0112591, 0.0908695]	Excludes 0

Note: N = 6014; 821 firm clusters; 1000 bootstrap replications. a is the policy-to-mediator coefficient, b is the mediator coefficient conditional on DID, and ab is their product. Values retain seven decimal places. Both types of 95% confidence interval are reported.

The two products should not be added to construct a total mediated share because the channels are estimated separately and may overlap. Nor does either interval resolve the additional identification assumptions required for causal mediation. These findings provide mechanism-consistent evidence rather than a definitive decomposition of the policy's causal effect.

4.6 Ownership Heterogeneity

Table 10 compares ownership groups. The DID coefficient is 0.2645 for non-state-owned enterprises and is significant at the 1% level. For state-owned enterprises, it is 0.0980 and is not significant at conventional levels. The larger point estimate for non-state-owned firms is compatible with a higher marginal value of liquidity where financing frictions are more pronounced. However, the table does not include a formal test of the cross-group coefficient difference, so it cannot establish statistically significant ownership heterogeneity.

Table 10: Estimates by ownership

Variable	Full sample	SOEs	Non-SOEs
DID	0.2202***	0.0980	0.2645***
	(0.0592)	(0.1204)	(0.0672)
Rebates	22.6524***	24.3199***	21.6491***
	(2.0445)	(5.8375)	(2.1714)
Size	0.8549***	0.7819***	0.8584***
	(0.0546)	(0.1280)	(0.0626)
Lev	0.1693	-0.1447	0.1555
	(0.1793)	(0.5592)	(0.1981)
ROA	0.7739***	0.5858	0.7641***
	(0.2070)	(0.5189)	(0.2161)
Constant	-0.0878	1.9591	-0.2104
	(1.1908)	(2.7748)	(1.3573)
Firm/year fixed effects	Yes	Yes	Yes
Observations	6014	1153	4850
R^2	0.8966	0.9188	0.8902

*Note: Firm-clustered robust standard errors are in parentheses. All models include firm and year fixed effects. ***, **, and * indicate significance at the 1%, 5%, and 10% levels. Reported group-specific sample sizes are retained; no cross-group coefficient-equality test is reported.*

5. Conclusion

This study links VAT credit refund exposure to firms' overseas operating performance using an unbalanced panel of 6014 observations from 821 Chinese listed firms over 2013–2023. The fully controlled DID coefficient is 0.2202 log points and is significant at the 1% level. The event-study test does not reject parallel

© 2026 The Author(s). Authors retain copyright in their articles.

pre-policy trends, and the placebo, matching, weighting, and reduced-period estimates are consistent with the positive baseline result. A complementary regression also finds a positive relationship with firm resilience.

The mechanism results connect the policy with lower financing constraints and higher total factor productivity. Firm-clustered bootstrap intervals exclude zero for both estimated indirect effects. These patterns complement the evidence on financial barriers to export participation and VAT credit refunds as a source of financing relief. They also relate to research on VAT export rebates and upgrading, while examining a different refund instrument and a broader overseas-revenue outcome. The contribution is therefore evidence on financial and productive conditions supporting export development, rather than direct identification of product-quality improvements.

The policy implications concern both liquidity provision and the productive use of released funds. Predictable and timely refunds can support working capital planning and reduce financing pressure. Complementary services that help firms evaluate investment projects, improve production processes, and manage overseas operations may strengthen the link between short-term liquidity and longer-term capability. The larger point estimate among non-state-owned firms suggests that access barriers deserve attention, although the present results do not justify claiming a statistically established ownership difference.

Several limitations define the scope of the findings. First, overseas operating revenue is a scale-based outcome, and future work should use customs-linked product quality, export prices, domestic value added, or export survival measures. Second, the long sample spans later refund expansions and other changes in the trading environment; alternative exposure definitions and policy-specific windows would strengthen identification. Current-year industry coding should also be compared with a treatment group fixed before the reform. Third, sparse early observations, industry-level dependence, and potentially policy-responsive controls warrant further sensitivity checks. Finally, the mechanism evidence requires additional assumptions for causal interpretation, while ownership differences require a direct statistical test. These extensions would clarify how far improved overseas performance translates into high-quality export development.

References

[1] Ceglowski J. Assessing export competitiveness through the lens of value added. The World Economy. 2017; 40(2):275-296.

[2] Bellone F, Musso P, Nesta L, Schiavo S. Financial constraints and firm export behaviour. The World Economy. 2010; 33(3):347-373.

[3] Greenaway D, Guariglia A, Kneller R. Financial factors and exporting decisions. Journal of International Economics. 2007; 73(2):377-395.

[4] Pernet T. Value-added export tax policy, credit constraints, and quality: Evidence from China. Review of Development Economics. 2024; 28(2):499-526.

[5] Yue S, Xiao C. Can VAT credit refunds alleviate corporate financing constraints? Empirical evidence based on cash–cash flow sensitivity. Finance & Trade Economics. 2023; 44(1):51-67. [In Chinese].

[6] Ministry of Finance, State Taxation Administration. Notice on tax policies concerning the refund of VAT credits in selected industries in 2018 (Cai Shui [2018] No. 70). 2018. Available from: https://www.chinatax.gov.cn/chinatax/n810341/n810765/n3359382/201807/c3730844/content.html [Accessed 11 September 2026].

[7] State Taxation Administration. Trial implementation of the VAT credit refund system from April 1: Eligibility requirements. 2019. Available from: https://www.chinatax.gov.cn/n810356/n3255681/c4177366/content.html [Accessed 11 September 2026].

[8] State Taxation Administration. Explanation of the 2022 VAT credit refund policy. 2022. Available from: https://www.chinatax.gov.cn/chinatax/c102191/c5177225/content.html [Accessed 11 September 2026].

[9] Whited T M, Wu G. Financial constraints risk. The Review of Financial Studies. 2006; 19(2):531-559.

[10] Levinsohn J, Petrin A. Estimating production functions using inputs to control for unobservables. The Review of Economic Studies. 2003; 70(2):317-341.

© 2026 The Author(s). Authors retain copyright in their articles.

[11] Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? The Quarterly Journal of Economics. 2004; 119(1):249-275.

[12] Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. Statistical Science. 2010; 25(1):51-71.

[13] Rosenbaum P R, Rubin D B. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.129142
Published by: Zeus Press

The Study of the Impact Mechanism of R&D Expense Super-deduction on Green Innovation of High-tech Enterprises in the Context of the Digital Economy

Ting Wang*

School of Public Finance and Taxation, Southwestern University of Finance and Economics, Chengdu 610000, China

**Corresponding author: Ting Wang.*

Abstract

As the digital economy improves by leaps and bounds, new factor conditions are provided for corporate green innovation. In the meantime, the R&D expense super-deduction policy has become a vital fiscal and tax instrument for incentivizing high-tech enterprises to conduct low-carbon technology R&D. Nevertheless, existing research rarely places the digital economy, the super-deduction policy, as well as corporate green innovation within the same framework, and analyses of transmission paths and differentiated policy effects still need improvement. Against this background, this research selects panel data of A-share listed high-tech enterprises from 2019 to 2023 as the sample, and then constructs two-way fixed effects and mediation effect models, and empirically examines the impact of the R&D expense super-deduction on green innovation of high-tech enterprises and its internal transmission mechanism. Further, Heterogeneity analysis is conducted from three dimensions: ownership nature, regional distribution, and firm size. The results suggest that the R&D expense super-deduction reinforces green innovation in high-tech enterprises, R&D intensity plays a partial mediating role in this process, and the policy incentive effect is heterogeneous, being more significant in non-state-owned, eastern-region, and small-scale high-tech enterprises. This paper supplements micro-level evidence on the intersection of fiscal and tax incentives and corporate green innovation, and provides references for the government, in order to implement differentiated tax policies and for enterprises to conduct green R&D.

Keywords

digital economy, R&D expense super-deduction, high-tech enterprises, green innovation, R&D intensity

1. Introduction

With the perspective of high-quality development and the “dual carbon” goal, the digital economy is undergoing quick iteration, and the R&D expense super-deduction policy has been put into effect in full by high-tech companies, which has become a crucial fiscal and taxation measure to ease business R&D tax loads, as well as promote green and low-carbon technologies. Both digital technologies and tax policies have an impact on the process of corporate innovation, however, in reality, the outcomes of policy implementation are not uniform.

© 2026 The Author(s). Authors retain copyright in their articles.

According to the available statistics and the information in the CSMAR database, the size of R&D spending in high-tech industries has increased over time, but the share of green patents in total patents is still small, and innovation assets are located in the east, whereas low-carbon R&D of high-tech companies in the middle and west are somewhat weak. Due to the long duration and high risk involved in green innovation, businesses tend to be financially constrained. Following the benefit of super deduction, there are some companies that have not used tax saving money into green R&D and digitalization has not been integrated with low carbon R&D effectively.

In the area of three fundamental facets of R&D expenses super-deduction, digital economy, and green innovations in business, both local and foreign scholars have developed a fairly abundant scholarly literature. Under the dual carbon objectives and policy coverage growth, more and more domestic scholars are starting to focus on the real incentives of policies to green innovations.

When it comes to the study of tax incentives and corporate innovation, the current literature has shown that R&D tax incentives are capable of encouraging businesses to invest more in R&D by removing financing constraints and freeing up internal cash flows; consequently, it would have an impact on the level of innovation output [1,2]. Other studies have also performed mediation and heterogeneity tests in which they examine the variations in policy effects based on factors like ownership and the corporate life cycle. Nevertheless, much of the empirical literature simply measures the overall number of corporate patents as an indicator of innovation level without any distinction between green and low-carbon patents, and it becomes hard to identify the particular influence of tax policies on green innovation [3]. A number of studies develop indicators based on combining various forms of tax incentive testing, but only a small portion focuses on examining the impact of the R&D expense super-deduction in a separate manner as a special policy towards green innovation [4]. Only a few studies use case analysis to examine the effects of policy implementation, however, due to the single-sample limitation of studies, their findings cannot be generalized. Early works, limited in time by the sample size, were unable to adequately consider the realistic context of the high rate of growth of the digital economy [5]. The practical literature is primarily concerned with institutional issues, including R&D cost collection and ledger management, and it generally takes the form of qualitative discussion, with no large-sample micro-level empirical work being conducted on the topic of green innovation [6,7].

One of the other significant areas of research in this regard is the impact of the digital economy on corporate innovation and green innovation. Previous studies have shown that at the urban macro level, the digital economy could optimize the distribution of regional innovation factors and enhance the pattern of spatial distribution of innovation resources [8]. Other sources, based on the point of view of digital finance, claim that digitalization may help to ease information asymmetry, minimize external financing expenses, and contribute to the empowerment of corporate R&D efforts indirectly [9]. There are some studies that suggest that digitalization also has two-fold effects of reducing costs and investing heavily, i.e., the digital innovation paradox [10]. Nevertheless, the current body of research has its shortcomings: most works consider the topic of the digital economy and fiscal and tax policies individually, but few of them explore how the variations in the levels of digital economy development across regions affect the incentive effects of taxes on innovations; many studies are still conducted at the urban macro level with no attempts made to test the mediating transmission mechanisms at the firm micro level [11].

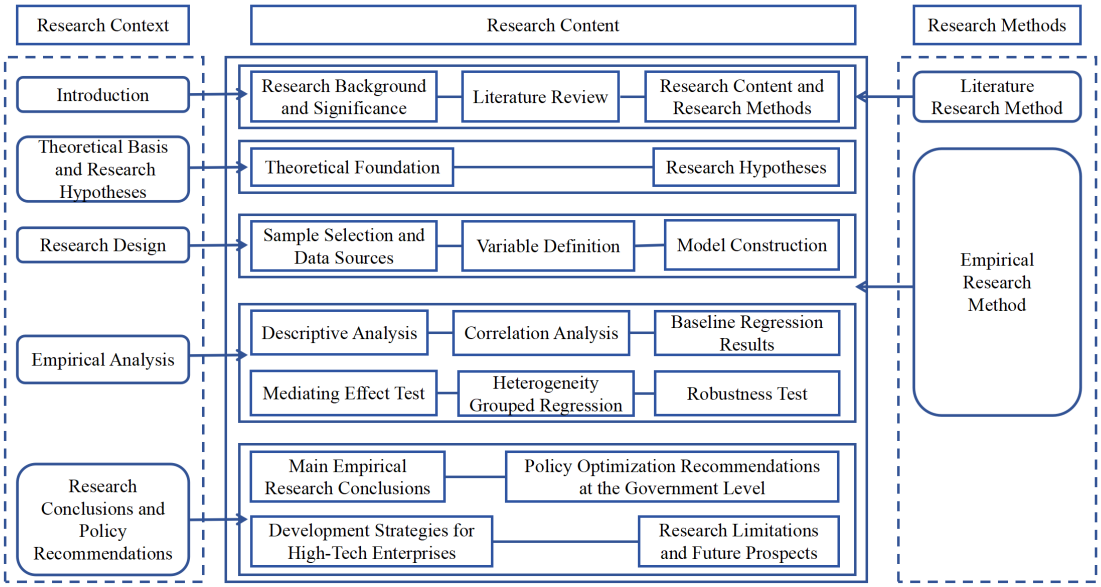
An overall review of all domestic and foreign literature shows that three main research gaps exist in the current body of work. The paper in particular contributes to the improvement of the current analytical framework so as to correct such shortcomings. Firstly, most literature uses total corporate patents to measure innovation output in general terms, but empirical evidence that specifically employs green patents to differentiate low-carbon innovation and develops fine-grained econometric models of A-share listed high-tech companies is rare. To pin down green innovation behavior accurately, this paper uses the logarithm of green patent applications as the independent variable. Secondly, the current studies divide the macro environment of the digital economy, the fiscal and tax policy of R&D super-deduction, and corporate green innovation, without testing the mediating transmission chain completely. This paper brings R&D intensity as an intervening variable and applies the three-step approach by Wen Zhonglin to systematically test the transmission process of R&D expense super-deduction-R&D intensity-green innovation, explaining in detail the inner mechanism through which the policy impacts green innovation and addressing the discontinuity of previous mechanism analysis. Thirdly, the presence of heterogeneity tests in the existing literature is disjointed and singular and lacks a coherent layered interpretation of differentiated policy impact incorporating both the dual aspects of

© 2026 The Author(s). Authors retain copyright in their articles.

ownership and region and does not explain the reason behind heterogeneity in light of the reality of the division of the domestic digital economy on the basis of region. This paper eliminates tests dimensions with no significant statistical support and keeps three groups of grouped regressions: ownership, region, and firm size. It examines differences in policy responses between state-owned and private enterprises, eastern and central-western regions, and large and small enterprises, and offers explanations of the results of grouped regressions based on regional disparities in the development of digital economies, thus enhancing the layered research framework of fiscal and tax incentives to green innovation. On the theoretical front, it could add to the body of micro-level empirical evidence in the area of fiscal and tax incentives and corporate green innovation. On the practical front, the empirical findings of this paper may not only serve as empirical guidelines to the government in terms of formulating layered and differentiated tax policies toward green innovation, but also assist high-tech enterprises in making more effective use of the super-deduction policy in conducting green R&D, which will allow advancing the dual carbon objectives and industrial digitalization simultaneously.

Moreover, preliminary non-empirical findings indicate that the digital economy interaction coefficient is insignificant. No moderating effect model is proposed in the paper, since the digital economy is only included in the theoretical framework as part of the macro context, which corresponds to the objective features of the sample data and is more rigorous.

Figure 1: Research Framework Diagram



The paper is essentially based on two research approaches. The first approach is known as the literature review method whereby related domestic and foreign literature is reviewed, the available research findings and limitations are summarized and the theoretical basis of the paper is developed. The other approach is empirical research where panel data of A-share high-tech firms are chosen, and two-way fixed effects and mediating effect models are created and empirical analysis is done using Stata software. At the same time, heterogeneous regressions are performed to contrast how policies affect various ownership types, areas and company sizes.

2. Theoretical Basis and Research Hypotheses

2.1 Tax Incentive Theory

The tax incentive theory indicates that green innovation has important positive externalities. If the market were purely competitive, then business low-carbon R&D spending would be less than the socially optimal amount, causing market failure. To address this discrepancy the government can use tax-base preference tools to alter the cost-benefit structure of corporate R&D. The R&D spending super-deduction policy is an average tax-base preference that permits businesses to make additional deductions of R&D expenditures prior to taxation with a particular rate. The mechanism is reducing the marginal cost of green innovation directly,

© 2026 The Author(s). Authors retain copyright in their articles.

reducing the enterprise tax burden on low-carbon R&D, and thus fundamentally changing the enterprise attitudes towards conducting green patent R&D.

In line with the theoretical reasoning stated above, it can be concluded that the greater the extent of the super-deduction preference received by an enterprise, the more abundant are the funds accessible to low-carbon technology R&D, and the higher the level of green innovation output correspondingly. Meanwhile, the digital economy can open up cross-regional circulation channels for innovation factors, which in turn will lower the trial-and-error costs of low-carbon R&D and become an external environmental condition for enterprises conducting green innovation. Therefore, this paper suggests the research hypothesis of the main effect:

H1: The amount of R&D expense super-deduction has a significant positive promoting effect on green innovation of high-tech enterprises.

2.2 Financing Constraint Theory

Green innovation is characterized by a protracted R&D process and a greater level of information asymmetry. It is subject to more stringent internal and external financing conditions than traditional innovation. Poor internal cash flow, as well as external green credit premiums, will have a direct impact on the inability of low-carbon R&D activities to take place. The R&D expense super-deduction would relieve financial challenges along two routes: firstly, by decreasing corporate income tax spending and increasing internal operating cash flow; secondly, the history of enterprises that benefit from the tax preference will send a message of high-quality R&D to the capital market, reducing the barrier to external financing. Nevertheless, the additional funds created by tax savings cannot be immediately turned into green patent production, they need to go through an intermediary step of investing in R&D resources in order to make the innovation results effectively implemented. This transmission logic is used as a theoretical framework behind the further mediating effect test.

2.3 R&D Investment Mediation Theory

Mediation theory of R&D investment suggests that different fiscal, tax, and industrial policies cannot directly produce innovation output. R&D resource investment intensity is one major intermediate vehicle for converting policy dividends into innovation results. R&D expenditure as a proportion of operating revenue may be used as an objective indicator of the extent to which resources are inclined towards innovation efforts within an enterprise. Together with the capital transmission mechanism underlying the theory of financial constraints, the cash flow generated by the super-deduction policy will initially increase the total R&D budget of the enterprise and thus increase R&D intensity. Adequate R&D resources will keep on enabling low-carbon process development and environmental protection product updates until they eventually become green patent output. This creates a full chain of transmission between fiscal and tax policy and R&D investment leading to green innovation, which is also the main theoretical framework of the mediation logic in this paper. Thereby, the next hypothesis is formulated:

H2: R&D intensity plays a mediating role between the R&D expense super-deduction and green innovation of high-tech enterprises.

2.4 Enterprise Heterogeneity Theory

The theory of enterprise heterogeneity implies that various enterprises have organized differences in terms of ownership goals, location assets, and size of operations and will result in varied output impacts when the same fiscal and tax policies are applied. In regard to the ownership perspective, private companies consider market outcomes as their core operating aim, and green innovation is one of the effective approaches to the development of market competition, whereas state-owned companies have to strike a balance between various non-profit goals, including labor safety and supply chain stability, and therefore there are not many strict limitations on low-carbon R&D evaluation. In terms of the regional aspect, eastern cities have a strong presence of digital infrastructure, green industrial chains, and advanced scientific and technological innovation talent, whereas the central and western regions have significant deficiencies in digital auxiliary devices and innovation components. As per the classification of the National Bureau of Statistics, this paper has classified the regions to which the enterprises belong into the eastern region (13 provinces and municipalities): Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan, Liaoning, Jilin, and

© 2026 The Author(s). Authors retain copyright in their articles.

Heilongjiang, and the central and western regions (the rest of the provinces). The eastern region has a high concentration of the digital economy and scientific and technological innovation factors, and the central and western regions have severe weaknesses in innovation support facilities; it is anticipated that the policy effects will differ by region. On the scale dimension, the aggregate amount of R&D resources of large-scale enterprises is very large and the marginal amount of tax reduction is small, whereas the R&D base of small-scale enterprises is relatively low and the same tax-saving resources can be more effective in making them invest in low-carbon activities. According to these three types of endowment differences, the layered heterogeneity model of policy incentive effects can be inferred. Therefore, the hypotheses below are formulated:

H3: Compared with state-owned enterprises, the R&D expense super-deduction has a more significant promoting effect on green innovation of non-state-owned enterprises.

H4: Compared to companies in the central and western parts of the country, the R&D expenditure super-deduction is more influential in promoting green innovation of companies in the eastern part of the country.

H5: Compared to large-scale enterprises, R&D expense super deduction is a much more encouraging factor in green innovation of small-scale enterprises.

3. Research Design

3.1 Sample Selection and Data Sources

The present paper chooses listed companies that are identified as high-tech companies on the Shanghai and Shenzhen A-share markets during 2019-23 as the starting research sample. The screening of the sample is based on the principles of mainstream domestic empirical research: financial and real estate sectors are not included in the screening process; enterprises with financial irregularities (ST and *ST) are not included; and missing core variables are excluded with significant missing data.

Following repeated screenings and proper panel balancing, there is the final sample of 14,246 firm-year panel observations. The CSMAR (China Stock Market & Accounting Research) database provides firm-level micro financial indicators, green patent application data, R&D expense super-deduction amounts, and R&D expenditure data. The empirical regression coefficients are not affected by extreme outliers as two-sided winsorization is used in all continuous variables at 1% and 99%. The data processing and regression analysis were done using Stata 17.0 econometric software.

3.2 Variable Definition

3.2.1 Dependent Variable

The measure is the natural logarithm of the sum of the total number of green patent applications in the present year, including green invention patents and green utility model patents. The calculation formula is $LnGrePatent_{it}=ln(GrePatent_{it}+1)$.

The number of patent applications is taken instead of granted patents in this paper to prevent the problem of approval time lag; green patents are differentiated so as to be relevant to the topic of research in this paper; adding one and taking a logarithm solves the problem of too many values with zero patents in the sample and reduces right-skewness in the data.

3.2.2 Independent Variable

The paper picks the practical value of R&D expense super-deduction tax relief that the enterprise is using in a given year as the main independent variable, which is actually assessed as the ratio between the R&D super-deduction tax relief and the nominal income tax rate of the enterprise.

This continuous indicator may indicate the real magnitude of policy preference achieved by the enterprise. It is more sophisticated than a 0-1 dummy variable. The information is based on the income tax schedule of the CSMAR database, which has a uniform measurement scale and is able to evaluate the decrease in R&D expenses due to the policy.

© 2026 The Author(s). Authors retain copyright in their articles.

3.2.3 Mediating Variable

The measure is the ratio of enterprise R&D expenditure to operating revenue, the intensity of overall R&D resource investment by the enterprise.

The measure of R&D intensity refers to the division of R&D expenditure by operating revenue. It could remove the influence of the scale of a firm and enable the horizontal comparison of the investment in R&D resources of enterprises. The variable constitutes the central part of the mediating path in the present paper. Equivalent domestic research on mediating mechanisms usually uses this indicator [12] with enough literature support and makes sure that results of empirical investigations are consistent and reproducible.

3.2.4 Control Variables

With reference to the existing literature, this paper chooses the following variables: size of the firm (*Lnsize*), age of the firm (*LnAge*), ratio between assets and liabilities (*Lev*) and operating cash flow (*CFS*). Besides, it controls for both the firm-specific fixed effects and year fixed effects simultaneously to minimize the effect of omitted variables.

Table 1: Variable Definitions

Category	Variable Name	Variable Definition
Dependent Variable	*LnGrePatent*	ln (Number of green patent applications + 1)
Independent Variable	*Deductamt*	Tax relief amount from R&D expense super-deduction / Enterprise's nominal income tax rate
Mediating Variable	*RD*	R&D expenditure / Operating revenue
Control Variables	*Lnsize*	ln (Total assets + 1)
	LnAge	ln (Observation year − Year of firm establishment + 1)
	Lev	Total liabilities / Total assets
	CFS	Net cash flow from operating activities / Total operating revenue

3.3 Model Construction

3.3.1 Baseline Regression Model

In order to test the main effect hypothesis H1, the following two-way fixed effects panel baseline regression model is designed:

$$LnGrePatent_{it}=\alpha_0+\alpha_1 Deductamt_{it}+\alpha_2 Controls_{it}+\mu_i+\lambda_t+\varepsilon_{it} \quad (1)$$

In which, $LnGrePatent_{it}$ is the green innovation output level of the i-th high-tech enterprise at the end of the period t; $Deductamt_{it}$ is the amount of R&D expense super-deduction of an enterprise in the current period; $Controls_{it}$ is the set of all control variables; μ_i is the firm individual fixed effect; λ_t is the year time fixed effect; ε_{it} is the random disturbance term of the model. The support of Hypothesis H1 is given when the key coefficient α_1 is significantly larger than 0.

3.3.2 Mediating Effect Model

Based on the stepwise regression test method of Wen Zhonglin and Ye Baojuan [13] three progressive regression model sets are formed:

Model 1 (Baseline Total Effect Regression):

$$LnGrePatent_{it}=\alpha_0+\alpha_1 Deductamt_{it}+\alpha_2 Controls_{it}+\mu_i+\lambda_t+\varepsilon_{it} \quad (1)$$

Model 2 (Independent Variable Regressed on Mediating Variable):

$$RD_{it}=\beta_0+\beta_1 Deductamt_{it}+\beta_2 Controls_{it}+\mu_i+\lambda_t+\varepsilon_{it} \quad (2)$$

Model 3 (Simultaneously Incorporating Independent Variable and Mediating Variable):

$$LnGrePatent_{it}=\gamma_0+\gamma_1 Deductamt_{it}+\gamma_2 RD_{it}+\gamma_3 Controls_{it}+\mu_i+\lambda_t+\varepsilon_{it} \quad (3)$$

When the coefficients α1 in Model 1, β1 in Model 2 and γ2 in Model 3 are statistically significant, and the coefficient γ1 of the core independent variable in Model 3 decreases in value or becomes less significant than

© 2026 The Author(s). Authors retain copyright in their articles.

α1 in Model 1, then R&D intensity acts as a partially mediating transmission factor, and Hypothesis H2 is confirmed.

3.3.3 Heterogeneity Grouped Regression Model

The three subsamples are sorted by the type of ownership, location of registration, and industry median of total assets, and individual regression models are fitted to test hypotheses H3, H4 and H5.

4. Empirical Analysis

4.1 Descriptive Statistics

Table 2: Descriptive Statistics Results

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
LnGrePatent	14246	0.568	0.940	0	4.025
Deductamt	14246	7.12e+07	1.50e+08	0	1.01e+09
RD	14246	7.5288	7.229	0.410	44.960
Lnsize	14246	8.175	1.108	6.079	11.581
LnAge	14246	3.005	0.283	2.197	3.611
Lev	14246	0.370	0.191	0.049	0.865
CFS	14246	0.099	0.158	-0.473	0.569

Descriptive statistics findings of all variables are presented in Table 2. The analysis indicates that the output of green innovation in high-tech enterprises is distinctly differentiated in the sample. Over half of the enterprises did not apply for any green patents this year, which suggests inadequate overall green innovation enthusiasm and modest results, as well as an uneven structure of low-carbon technology R&D. Such a feature corresponds to the result of critical disparity among the levels of enterprise innovation by Wang Huixin et al. (2025) [12] findings. Super-deduction preferences enjoyed by various enterprises differ significantly: small and medium high-tech enterprises have a small amount of R&D basis and a small size of tax savings, whereas the most prominent big enterprises can benefit from significant tax relief. High-tech enterprise internal R&D allocation is also not even, and many enterprises have invested less in R&D compared to the industry average. Control variable distribution is reasonable, and the sample represents high-tech companies of varying sizes and at various stages of operation, which is indicative of a high quality of sample representativeness.

4.2 Correlation Analysis

Table 3: Correlation analysis results

Variable	*LnGrePatent*
Deductamt	0.218^{***}
RD	0.0146^{*}
Lnsize	0.334^{***}
LnAge	0.0108
Lev	0.222^{***}
CFS	-0.0561^{***}
Observations	14246

Note: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The Pearson correlation outcomes indicate that the core explanatory variable has a significant positive relationship with the dependent variable as expected by the theory. All the correlation coefficient absolute values between any two variables are smaller than 0.4 which preliminarily confirms that there is no serious multicollinearity and that the variables can be used in regression analysis.

4.3 Baseline Regression Results

The reduction in the size of the baseline regression sample originates from the deletion of single-period firm observations through two-way fixed effects and the elimination of missing values on variables. The actual effective number of observations is 13,667, which does not alter the conclusions.

The value of the core explanatory variable, namely the R&D expense additional deduction amount (*Deductamt*) coefficient, is 0.000 and statistically significant at a 10 percent level, which means that the growth

© 2026 The Author(s). Authors retain copyright in their articles.

of the R&D expense additional deduction amount is positively correlated with the increase in the green innovation output of high-tech enterprises, which confirms H1.

Table 4: Baseline Regression Results

Variable	*LnGrePatent*
Deductamt	0.000*
	(0.000)
Lnsize	0.118***
	(0.027)
LnAge	0.453*
	(0.274)
Lev	0.041
	(0.079)
CFS	-0.085*
	(0.047)
Year fixed effects	Controlled
Individual fixed effects	Controlled
Constant	-1.665**
	(0.813)
Observations	13667

*Note: * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors are given in parentheses.*

The extra deduction reduces the marginal cost of corporate green R&D by lowering the tax base, adds more internal operating cash flow, relieves long-term and massive funding pressures on green innovation, encourages companies to invest in low-carbon process and environmental protection product R&D, and eventually brings up more green patent applications.

The coefficient of the control variable, firm size (*Lnsize*), is 0.118 and statistically significant at the 1% level, which implies that big firms are likely to be more successful in the establishment of a well-developed R&D system and having adequate capital resources; the coefficient of operating cash flow (*CFS*) is statistically negative, implying that companies with excess cash flow can invest their money into short-term lucrative opportunities, leaving less space to invest in green R&D with long payback periods and low yields; there are no statistically significant effects of firm age (*LnAge*) and asset-liability ratio (*Lev*).

4.4 Mediation Effect Test

In accordance with the three-stage mediation process of Wen Zhonglin, three successive regression models are built. Model (1) is the baseline total effect test; Model (2) takes R&D intensity (RD) as the dependent variable; and Model (3) takes both Deductamt and the mediator RD. The findings are given in Table 5.

Regression findings of Model 1 in Table 5 indicate that the R&D expense super-deduction amount coefficient (*Deductamt*) is statistically significantly positive, which means that the R&D expense super-deduction has a general promoting impact on corporate green innovation. In Model 2, R&D intensity (RD) is taken as the dependent variable, and the core explanatory variable, the R&D expense super-deduction amount (*Deductamt*), is highly significant at the 1 percent level. Tax relief savings could be used to directly increase corporate R&D expenditure and make the share of R&D spending in operating revenue significantly larger. This model includes both the core explanatory variable (*Deductamt*) and the mediator (RD). The R&D expense super-deduction amount (*Deductamt*) coefficient is slightly less than with Model 1, but its significance remains the same; the regression coefficient of RD is positive, which is consistent with the partial mediator criterion, thus Hypothesis H2 is confirmed. It can then be inferred that, apart from directly enhancing green innovation in high-tech companies, the R&D expense super-deduction also has an indirect transmission path: tax benefits lower corporate tax payments; companies use the extra money to fund R&D efforts, which leads to increased overall R&D intensity; adequate R&D capacity enables repetitive development of low-carbon technologies; and finally, the amount of green patent filings rises. R&D intensity partially mediates between these two.

© 2026 The Author(s). Authors retain copyright in their articles.

Table 5: Mediation Effect Test Results

Variable	*LnGrePatent* Total Effect	*RD* Mediator	*LnGrePatent* Direct Effect
Deductamt	0.000*	0.000***	0.000*
	(0.000)	(0.000)	(0.000)
RD	-	-	0.003*
			(0.002)
Lnsize	0.118***	-1.424***	0.122***
	(0.027)	(0.159)	(0.027)
LnAge	0.453*	10.460***	0.426
	(0.274)	(1.832)	(0.274)
Lev	0.041	-0.755	0.043
	(0.079)	(0.590)	(0.079)
CFS	-0.085*	-3.990***	-0.075
	(0.047)	(0.474)	(0.048)
Year/Individual fixed effects	Controlled	Controlled	Controlled
Constant	-1.665**	-11.306**	-1.635**
	(0.813)	(5.223)	(0.813)
Observations	13667	13667	13667

*Note: * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors in parentheses.*

4.5 Grouped Regression for Heterogeneity Analysis

In order to test two sets of heterogeneity hypotheses (H3, H4, and H5), the sample is split into three categories based on ownership nature, region of registration, and size of the firm to perform two-way fixed effects regressions between groups to determine how much the policy incentive effects differ among groups.

4.5.1 Heterogeneity Analysis by Ownership Nature

Table 6: Heterogeneity Test Results by Ownership Nature

Variable	*LnGrePatent* (State-owned Enterprises)	*LnGrePatent* (Non-state-owned Enterprises)
Deductamt	-0.000	0.000**
	(0.000)	(0.000)
Lnsize	0.075	0.120***
	(0.059)	(0.031)
LnAge	0.943	0.397
	(0.767)	(0.302)
Lev	0.147	-0.003
	(0.212)	(0.081)
CFS	-0.099	-0.098**
	(0.128)	(0.049)
Year/Individual fixed effects	Controlled	Controlled
Constant	-2.702	-1.534*
	(2.324)	(0.899)
Observations	2708	10959

*Note: * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors in parentheses.*

The research also separates the sample by state-owned enterprises (SOEs) and non-state-owned enterprises (non-SOEs) based on the characteristics of the actual controller. The grouped regression results are given in Table 6. These results indicate that the regression coefficient of the R&D expense super-deduction amount does not have a significant value in the SOE subsample, but it is significantly positive at the 5 percent level in the non-SOE subsample, thus confirming Hypothesis H3.

Non-SOEs compete under a highly competitive market setting and green innovation has a direct relationship to their market competitiveness. On the other hand, these companies are limited in their internal capital and have more stringent external finance restrictions, which means that the additional cash flow associated with tax relief may be used to invest in low-carbon R&D. By comparison, SOEs have diverse business goals. Besides revenues generated through operations, SOEs have to take on social mandates including employment and industry safety. They are less restricted in their performance evaluation of green R&D. Furthermore, SOEs are able to have greater access to credit facilities and experience lower financial pressures. The cash flow due

© 2026 The Author(s). Authors retain copyright in their articles.

to tax incentives is a marginal increase in the amount of money saved on taxes and will not significantly alter the initial course of corporate fund allocation. Thus, the super-deduction is not likely to increase their green patent production.

4.5.2 Regional Heterogeneity

Table 7: Regional Heterogeneity Test Results

Variable	*LnGrePatent* (Eastern Enterprises)	*LnGrePatent* (Central and Western Enterprises)
Deductamt	0.000**	-0.000
	(0.000)	(0.000)
Lnsize	0.124***	0.079*
	(0.032)	(0.046)
LnAge	0.491	0.382
	(0.311)	(0.585)
Lev	-0.003	0.148
	(0.088)	(0.153)
CFS	-0.088	-0.074
	(0.055)	(0.099)
Year/Individual fixed effects	Controlled	Controlled
Constant	-1.799*	-1.202
	(0.929)	(1.694)
Observations	10497	3170

*Note: * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors in parentheses.*

The sample is split into two categories by the location of the registered enterprise, which are the eastern region and the central and western regions. The outcomes of regional grouped regression analysis are given in Table 7. It is found that the coefficient of the main explanatory variable for high-tech enterprises in the eastern region is highly significant at the 5 percent level, but the regression coefficient of the same variable for enterprises in the central and western regions fails to be significant. Thus, Hypothesis H4 is supported by empirical evidence.

The identical tax policy causes disparate results in different regions as it is directly connected with regional innovation resources and the favourable environment of the digital economy. The eastern region has intensive digital infrastructure, green industrial chains, and leading scientific and technological innovation talent. Once companies have secured tax-saving money, they will be able to convert these assets into green patents through supportive R&D systems and industrial conditions. In comparison, the central and western regions have weaknesses in the provision of factors concerning green innovation. Although enterprises may get the same tax stimuli, they do not have digital R&D carriers, talent, and industrial chain support, so that the funds can easily be directed towards conventional technological iterations and are hard to convert into green innovation output. It also explains theoretically why the digital economy interaction term is insignificant in the sample: the influence of tax incentives is extremely sensitive to local real innovation conditions. The eastern region can provide policy dividends, whereas the central and western regions can balance the policy impact through the factor deficiencies. A single tax incentive can never make up for the systemic gap in regional innovation factors; both fiscal and tax policies and regional public supporting facilities should be synergized.

4.5.3 Heterogeneity by Enterprise Size

The sample is segmented into two categories, which are large-scale and small-scale companies, depending on the industry median of annual total assets. Table 8 contains the regression findings of the heterogeneity of enterprise sizes. As indicated by the results of the grouped regression by size, the regression coefficient of the R&D expense super-deduction amount does not matter statistically in large-scale enterprises, but the relevant coefficient in small-scale enterprises is highly significant at the 5 percent level. Thus, Hypothesis H5 is confirmed. This disparity has a fundamental cause, which is the varying marginal effects of tax-saving funds. In large-scale high-technology enterprises, there exists a large R&D fund pool of overall spending, and the savings brought by the super deduction represent a relatively small fraction of total R&D investment, therefore, incremental funds play a minimal role in the overall R&D plan. Moreover, large companies operate in diverse fields, and tax-saving gains do not necessarily go towards green R&D but rather spread across various lines of production and operations. On the other hand, small high-tech enterprises have a lower R&D investment base,

© 2026 The Author(s). Authors retain copyright in their articles.

and the same amount of tax relief may result in much larger increases in the share of R&D investment. Meanwhile, small and medium businesses have constrained sources of financing, and tax benefits become an inexpensive way to finance internal R&D activities. They are more responsive to policy, and tax-saving funds could be used to achieve green patent output successfully.

Table 8: Heterogeneity Test Results by Enterprise Size

Variable	*LnGrePatent (Large-scale Enterprises)*	*LnGrePatent (Small-scale Enterprises)*
Deductamt	0.000	0.000**
	(0.000)	(0.000)
Lnsize	0.158***	0.080**
	(0.044)	(0.039)
LnAge	0.637	0.349
	(0.466)	(0.360)
Lev	0.160	0.006
	(0.139)	(0.107)
CFS	-0.086	-0.023
	(0.081)	(0.061)
Year/Individual fixed effects	Controlled	Controlled
Constant	-2.529*	-1.180
	(1.387)	(1.051)
Observations	6929	6738

*Note: * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors in parentheses.*

4.6 Robustness Test

Table 9: Robustness Test Results

Variable	*LnGrePatent*
L_Deductamt_w	0.000*
	(0.000)
Lnsize	0.107***
	(0.033)
LnAge	0.106
	(0.354)
Lev	-0.061
	(0.095)
CFS	-0.080
	(0.054)
Year/Individual fixed effects	Controlled
Constant	-0.481
	(1.059)
Observations	10518

*Note: * p < 0.1, ** p < 0.05, *** p < 0.01; standard errors in parentheses.*

Due to the presence of reverse causality as a possibility, this paper regresses the primary explanatory variable with a one-period lag. The findings in Table 9 indicate that the coefficient of the lagged term is still statistically significantly positive at 10 percent, and the sign is the same. It means that the main conclusion is strong even after removing the influence of bidirectional causality. Funds to tax policies have a time lag, and even though the amount saved on taxes in any given year will continue to affect R&D in future years, this is also supported by the regression findings.

5. Research Conclusions and Policy Recommendations

5.1 Main Empirical Research Conclusions

The paper uses 14,246 firm-years of panel data of A-share listed high-tech companies for the period between 2019 and 2023 and runs the entire battery of empirical tests. Two-way fixed-effects baseline regression confirms that R&D expense super deduction will help to boost corporate green innovation output significantly, which means, H1 is corroborated. The three-stage mediation test of Wen Zhonglin indicates that R&D intensity has a mediating effect, which generates a full transmission process of the form R&D expense

© 2026 The Author(s). Authors retain copyright in their articles.

super deduction-increase in R&D intensity-green innovation promotion, and thus H2 is supported. The regroupings into ownership, regions and sizes show strong stratified differences in policy incentives. Just non-state-owned enterprises, eastern region enterprises and small scale high-tech enterprises can efficiently release policy dividends; the incentive impact on state-owned enterprises, central and western regions enterprises and large enterprises is insignificant. All these hypotheses, H3, H4 and H5, are empirically confirmed. The findings of the robustness test based on the one-period lagged core variable are also similar to the baseline regression, and they indicate that the main findings of the present study cannot be altered by the endogeneity issue of bidirectional causality and that the findings are reliable and credible.

5.2 Government-Level Policy Optimization Recommendations

According to the baseline regression, mediation effect test and heterogeneity findings depending on ownership, region, and size, the paper suggests differentiated fiscal and tax policy and optimisation of the use of public services recommendations.

Concerning the ownership aspect, it is possible to integrate green R&D investment and green patent output into the performance evaluation system of state-owned enterprises, as the policy incentive impact on state-owned enterprises was minimal; such integration will direct tax-saving resources to green innovations. On top of that, the procedures for collecting and declaring R&D expenses of privately owned and foreign-invested companies need to be constantly optimized, which will lower the institutional transaction costs and completely unleash the green innovation energy of the market-oriented entities.

On the regional scale, it is necessary to provide proper green R&D super deduction preferences to the high-tech companies in the middle and western regions, which should be backed up by low-carbon digital government services and enhanced efforts to attract science-based and innovative talent to make up the gap in regional innovation determinants. In the eastern region, it may use its strength in the digital industry to concentrate on incentives for high-value green invention patent production and further encourage frontier low-carbon technology studies.

At the level of the size of firms, there is a need to increase the specific tax incentives for the green R&D of small enterprises. In big high-tech companies, companies should be encouraged to enhance restrictions on green R&D funding investments and refrain from diverting too much of the tax dividends into regular business operations.

Based on the transmission mechanism, tax and science and technology departments could work together to enhance the detailed regulations of independent green R&D accounting, to facilitate the efficient conversion of super deduction benefits into enterprise R&D spending, to create low-carbon digital carriers of public services, to introduce green credit products based on tax credits, to relieve financial constraints, and to release the transmission channel of the R&D expense super deduction-R&D intensity-green innovation.

5.3 Development Countermeasures for High-Tech Enterprises

The policy dividends of the R&D expense super deduction should be well utilized by high-tech enterprises, and the internal path between policy dividends and green innovation transformation should be opened up. Companies may form dedicated funds to green R&D, have separate control over the tax savings created by the super deduction, and increase the pace of growth in green patents by increasing the level of R&D.

Lightweight low-carbon process improvements can be undertaken by small-scale enterprises depending on their available resources, whereas large-scale enterprises may set up specialized green R&D departments, define the amount of low-carbon R&D spending, and decrease the degree of fund scattering. State-owned enterprises must enhance internal green R&D incentives and reinforce internal evaluation of innovations. Regional digital innovation platforms should be actively connected to private and foreign-funded enterprises in order to enhance efficiency in R&D transformation.

Eastern regional enterprises may also depend on the local digital industry chains to create a synergistic effect between digitalization and green R&D. The central and western enterprises are advised to interact with the government measures to digitalize and support talents to compensate the inadequacy of innovation resources. All types of enterprises must have standardized green R&D financial registers, correct classification

© 2026 The Author(s). Authors retain copyright in their articles.

and aggregation of R&D expenditures, as well as guarantee the authenticity and reliability of R&D expenditure accounting.

5.4 Research Limitations and Future Prospects

The present study has these limitations. Firstly, the sample is limited to high-tech companies listed in the A-share market, and the findings cannot be simply extrapolated to other unlisted science and innovation organizations. Secondly, only the overall quantity of green patents is considered, without splitting invention patents (utility model patents) to determine how the policy affected innovation quality. Thirdly, the digital economy is viewed solely as a macro background, and its sub-indicators are not broken down to test moderation. Thirdly, the endogeneity remedy relies on the lagged independent variable only, and no remedies like DID or instrumental variables are used. Future studies may increase the sample size to non-listed enterprises, differentiate green patent types to investigate the quality of innovation, break down digital economy sub-indicators, apply DID and instrumental variables to further control endogeneity, and add variables to the extension research, including environmental regulation and state grants.

References

[1] Qiqi Yan. (2025). Research on the impact of R&D expense deduction policy on the innovation ability of high-tech enterprises. Master's thesis. Chinese Academy of Fiscal Sciences.

[2] Shaowei Jia. (2025). A study on the impact of the R&D expense super-deduction on technological innovation of high-tech enterprises: An analysis from the perspective of enterprise life cycle. Master's thesis. Shanxi University of Finance and Economics.

[3] Zhaochang Ma. (2024). Research on the Impact of R&D Expense Deduction on the Innovation Output of High-tech Enterprises. Master's thesis. Shandong University of Finance and Economics.

[4] Wang, H., Yang, J., & Zhu, N. (2024). Does tax incentives matter to enterprises' green technology innovation? The mediating role on R&D investment. Sustainability, 16 (14), 5902. https://doi.org/10.3390/su16145902.

[5] Wansi He. (2022). A study on the impact of the R&D expense super-deduction on innovation of high-tech enterprises: A dual perspective of financing constraints and ownership attributes [Master's thesis]. Zhejiang University of Finance and Economics.

[6] Liling Yang. (2026). Research on financial risk management in the R&D expense super-deduction of enterprises. Modern Marketing, (1), 67–69.

[7] Zejian Yang. (2025). A study on strategies of high-tech enterprises in response to the R&D expense super-deduction policy. Management & Technology of SME, (19), 163–165.

[8] Tao Zhao, Zhi Zhang, & Shangkun Liang. (2020). Digital economy, entrepreneurship, and high-quality economic development: Empirical evidence from urban China. Journal of Management World, 36(10), 65–75.

[9] Kong, T., Sun, R., Sun, G., & Song, Y. (2022). Effects of digital finance on green innovation considering information asymmetry: An empirical study based on Chinese listed firms. Emerging Markets Finance and Trade, 58(15), 4399–4411. https://doi.org/10.1080/1540496X.2022.2083953.

[10] Ruizhi Pang, Dongge Liu. (2022). The paradox of digitalization and innovation: Does digitalization promote enterprise innovation? An explanation based on open innovation theory. South China Journal of Economics, (9), 97–117.

[11] Jiang, J. (2025). A configurational pathway analysis of digital economy and green technology innovation. Modern Economy, 16(2), 284–299.

[12] Huixin Wang & Ping Kuang. (2025). Can additional deduction of R&D expenses enhance the innovation vitality of high-tech enterprises? Science Technology and Industry, 25(14), 235–239.

© 2026 The Author(s). Authors retain copyright in their articles.

[13] Zhonglin Wen & Baojuan Ye. (2014). Analyses of mediating effects: The development of methods and models. Advances in Psychological Science, 22(5), 731–745. https://doi.org/10.3724/SP.J.1042.2014.00731.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.143161
Published by: Zeus Press

The Evolution of Predictive Analytics: Machine Learning and the Predictive Step in Empirical Asset Pricing—From Factor Zoo to Factor Refinery

Kaiyu Wu*

School of Business, Nanjing University of Science and Technology Zijin College, Nanjing 210023, China

Abstract

In the last fifty years, empirical asset pricing has depended on a consistent knowledge-production pipeline: economic theory first suggests hypotheses, hypotheses are then operationalised through candidate factors, and the validity of those hypotheses is finally determined by in-sample statistical significance. As the number of candidate factors grows into the hundreds, this pipeline is no longer a reliable way of identifying genuine factors among data-mined false positives. Organised around four steps — testing, construction, discovery, and integration — the paper argues that machine learning has established, alongside the conventional pricing tests, an evaluation framework based on out-of-sample predictive performance. This does not mean, however, that the out-of-sample evaluation framework is a substitute for alpha, GRS, or risk-price tests. The main thesis is that machine learning has made genuine instrumental advances in prediction-based tasks, but cannot be regarded as a substitute for causal inference or mechanistic explanation; its benefits are real but restricted, and its drawbacks (loss of interpretability, conflation of correlation with causation, and overfitting) are significant. A more precise statement is thus not that machine learning has changed empirical asset pricing, but that it has changed the mode of execution of prediction-related tasks within the field without changing the basic logic of understanding-related tasks.

Keywords

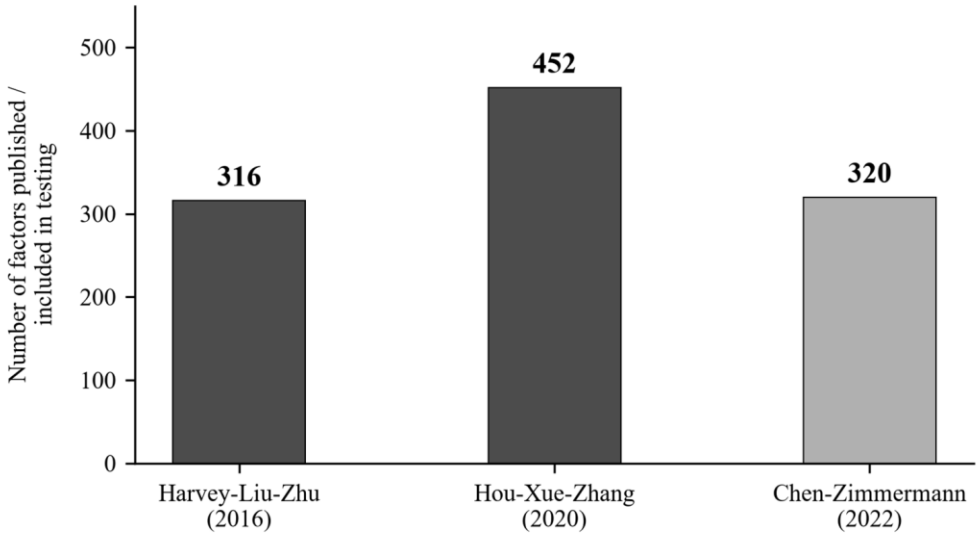
machine learning, asset pricing, factor zoo, out-of-sample prediction, evaluation framework

1. Introduction

The factor zoo is the honest self-assessment of the situation in empirical asset pricing. The term was introduced by Cochrane [1] in his discussion of discount rates; it implies that there are now so many potential factors that economic theory can no longer justify each of them individually. Harvey, Liu, and Zhu [2] reported 316 published cross-sectional return predictors; in the replication analysis of Hou, Xue, and Zhang [3], 452 anomalies were tested together; and the open-source factor data set of Chen and Zimmermann [4] lists more than 320 reproducible factors. The three counts differ in scale, but they all point in the same direction: the supply of factors passed the hundred mark long ago, far beyond what economic theory was previously expected to provide (see Figure 1, constructed and illustrated by the author).

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 1: The number of factors documented in three major studies: an expansion to several hundred



Compiled and drawn by the author from Harvey–Liu–Zhu 2016, Hou–Xue–Zhang 2020, and Chen–Zimmermann 2022. Note: The three counts are derived from different statistical bases, but they point in the same direction — the supply of factors far exceeds the a priori guidance suggested by economic theory.

Empirical asset pricing should reconsider the limits of its current testing structures now that "significance" is no longer believable and the replication requirements are not homogeneous. Machine learning presents exactly such an opportunity for re-examination, but its real value has long been misconstrued as the value of a more potent prediction tool. The paper proposes that the role of machine learning is not to replace the old pricing tests but to offer another benchmark of evaluation based on out-of-sample predictive performance, thereby compelling empirical asset pricing to reconsider what the validity of a factor means.

The argument proceeds in four steps — testing (Direction II, Section 3), construction (Direction III, Section 4), discovery (Direction IV, Section 5), and integration (Direction V, Section 6) — beginning with the issues raised by the conventional paradigm (Direction I, Section 2) and ending with its limits and the criticisms they invite (Direction VI, Section 7).

2. Challenges Facing the Traditional Paradigm

Conventional factor discovery has a well-defined structure: it starts with economic theory offering a hypothesis, continues with the construction of a candidate factor, and ends by testing the importance of that factor with data. Knowledge is generated along this path by theory alone; data play the role of arbiter. One of the main assertions of this section is that once the number of candidate factors rises into the hundreds, the conventional strategy of testing factors one by one and declaring them valid whenever they are significant becomes highly susceptible to scale pressure and can no longer separate genuine factors from false positives produced by data mining. This is not a failure of one particular study but a systemic problem of the paradigm under scale pressure.

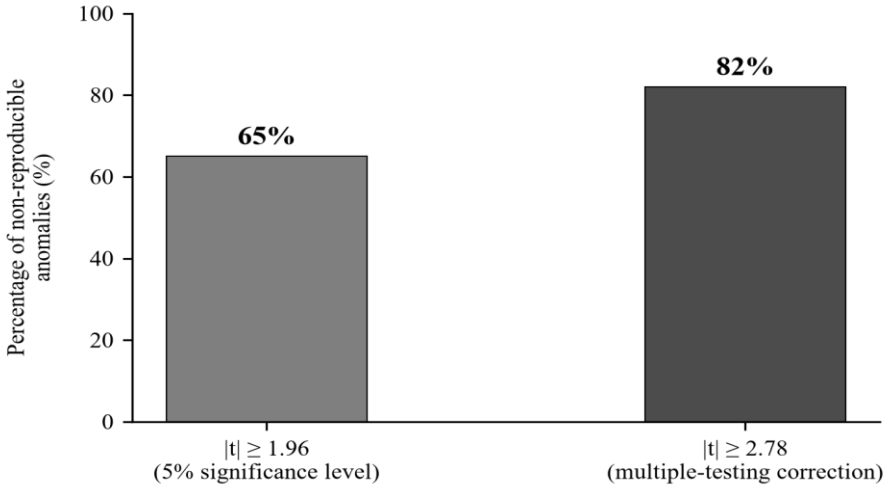
2.1 The Replication Crisis: Survival Rates under Unified Replication

The fundamental premise of the traditional paradigm is that when any element is considered important in an individual test, it is thus real. The assumption was directly tested using a so-called unified replication protocol by reexamination of 452 published anomalies with identical procedures, sample period, and control variables, as suggested by Hou, Xue, and Zhang [3].

The findings are clear: 65 percent of the anomalies were not reproduced at the significance level of $|t| \geq 1.96$; at the multiple-testing-adjusted level of $|t| \geq 2.78$, the failure rate rises to 82 percent (see Figure 2, redrawn by the author from Hou–Xue–Zhang 2020 data).

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 2: Anomaly failure rates under the unified replication protocol



Redrawn by the author from Hou–Xue–Zhang 2020 data. Note: $|t| \geq 1.96$ is the 5 percent significance level and $|t| \geq 2.78$ is the threshold for multiple-testing correction; the higher the threshold, the higher the failure rate, because much of the observed significance stems from multiple testing and data mining.

The two figures suggest that much of the earlier reported significance might have been due to multiple testing and data mining as opposed to actual economic effects.

However, a tension has to be recognised. Jensen, Kelly, and Pedersen [5], applying a Bayesian factor-correlation test to a global population of factors, report that most highly correlated factors are replicable, and therefore challenge the very name of a replication crisis.

These two studies are not contradictory; they differ in methodology. For Hou–Xue–Zhang, replication must be an exact re-run, that is, a test of whether the factor remains statistically significant under the same procedures as in the original study; for Jensen–Kelly–Pedersen, the methodology may vary and the question is whether the economic information behind the factor is still present. The first measures replicability, the second robustness. Since the interpretations of replication are not the same, the conclusions will also differ.

The value of this tension is that it reveals the underlying issue: the standard of replication in the traditional paradigm is not unique. What counts as valid depends on which replication standard is selected — and the traditional paradigm offers no unchallenged meta-standard for making that choice. This is why an evaluation model based on out-of-sample predictive performance is needed.

2.2 Three Supplementary Lines of Evidence: Multiple Testing, Marginal Contribution, and Publication Bias

Beyond the replication crisis, three independent chains of evidence point to the same conclusion.

The first is the multiple-testing problem. The fundamental thesis of Harvey, Liu, and Zhu [2] is a statistical truism that the classical paradigm long ignored: when each of hundreds of factors is tested separately, pure chance alone will produce a substantial number of so-called significant findings at the 5 percent level even if every factor is false. With more than 300 candidates, the conventional $t \geq 1.96$ limit is far too lax and needs to be tightened to $t > 3.0$. The argument is that the traditional pass criterion has already failed under an inflated number of factors, since it does not take into account how many factors were examined.

The second is the marginal contribution test. The more stringent question asked by Feng, Giglio, and Xiu [6] in “Taming the Factor Zoo” is whether a new factor adds any incremental information once the existing factors are controlled for. They conclude that most new factors are no longer incremental. The methodological importance of the work lies in its raising of the testing standard: the question is not simply whether a factor is significant, but whether it is incremental relative to current knowledge, which is much more demanding than the previous criterion.

© 2026 The Author(s). Authors retain copyright in their articles.

The third is the evidence on intertemporal publication bias. McLean and Pontiff [7] investigate the intertemporal performance of 97 variables and find that out-of-sample returns are 26 percent lower and post-publication returns 58 percent lower than in-sample returns. The decay is not random but systematic after publication, which can be read as evidence that the publication process has filtered out the best-performing statistical artefacts in sample.

2.3 Section Summary

The thesis of Direction I can be summarised in three statements. First, hundreds of factors have been identified, far more than economic theory could previously justify. Second, unified replication shows that 65–82 percent of anomalies do not replicate under a rigorous protocol, and the notion of replication itself is not unique in the conventional paradigm. Third, the joint indication of multiple testing, marginal contribution, and publication bias suggests that the traditional pass criterion is seriously threatened by scale pressure.

These three points establish a premise: the traditional paradigm faces systemic challenges under scale pressure. The premise provides the problem context for the expansion of the testing-evaluation framework presented in Section 3, but it should be noted that problems caused by scale pressure do not by themselves imply that a machine-learning framework can replace traditional tests.

3. Expanding the Testing-Evaluation Framework: From Significance Tests to Out-of-Sample Evaluation

Testing is the first step of empirical asset pricing and the one most directly affected by machine learning. The main argument of this part is that the obstacles facing conventional tests lie less in the severity of the significance levels than in structural flaws of the procedures themselves; machine learning has supplied an independent measure of factor assessment by introducing out-of-sample predictive performance. It should be noted that the machine-learning framework does not replace alpha, GRS, or risk-price tests but introduces, alongside the traditional tests, a benchmark evaluation based on out-of-sample predictive performance.

3.1 Inherent Limitations of the Traditional Testing Framework

Conventional tests rest on two classical tools whose logics differ and which should be considered separately.

The cross-sectional regression of Fama and MacBeth [8] first estimates the factor loadings β and then, in each cross section, uses those β to estimate the risk prices λ. It has a defect that has long been discussed: because β is itself estimated, it must contain measurement error, so when an investigator estimates λ from error-prone β, the classical errors-in-variables (EIV) problem emerges and distorts the estimate of λ. The flaw is partly hidden in portfolios but is greatly magnified in individual stocks.

The GRS test of Gibbons, Ross, and Shanken [9] is an alternative approach; in a time-series regression setting, it examines whether the alphas of all test assets are together zero (an F-test). It is not subject to the EIV critique described above because it does not require the step of estimating λ with β.

The criticism of the structure of test assets by Lewellen, Nagel, and Shanken [10] focuses on the fact that as long as test assets move in the same direction (which is almost unavoidable), nearly any factor mixture can appear to be reasonably priced. Put differently, the very criterion that cross-sectional tests must meet may be too lenient: a model need not have passed the test on its merits if the test itself was too simple to fail.

Jegadeesh et al. [11] respond by replacing single stocks with portfolios and by correcting the EIV bias of the Fama–MacBeth model with instrumental variables. The correction nevertheless remains within the bounds of the linear testing framework and does not touch the basis of the testing standard.

Note that the deficiencies mentioned above concern factor pricing only; they do not mean that the Fama–MacBeth and GRS techniques cannot be applied successfully in other contexts.

3.2 The Out-of-Sample Evaluation Benchmark of Machine Learning

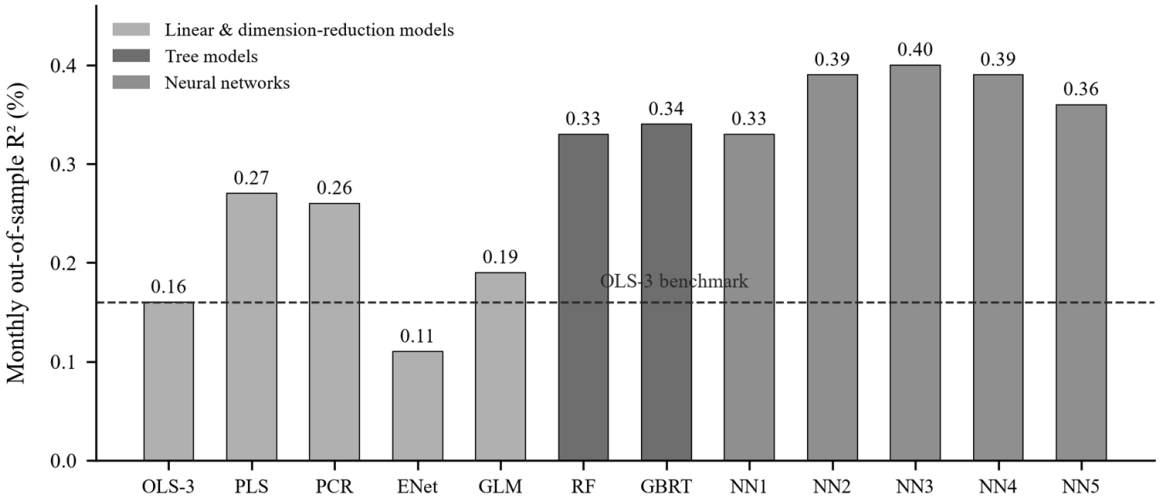
The main change brought about by machine learning is that factor validity is no longer identified with statistical significance in classical pricing tests but with strong predictive performance out of sample. This

© 2026 The Author(s). Authors retain copyright in their articles.

criterion adds a dimension to the definition of validity. It was systematically developed by Gu, Kelly, and Xiu [12].

Gu, Kelly, and Xiu [12] compare 13 prediction models, namely full-variable OLS, three-variable OLS, partial least squares (PLS), principal-component regression (PCR), elastic net (ENet), generalized linear models (GLM), random forests (RF), gradient-boosted regression trees (GBRT), and five neural networks (NN1–NN5). Their main contribution is the finding that nonparametric techniques, including tree models and neural networks, achieve the highest out-of-sample performance, with the best model reaching an out-of-sample R^2 of about 0.4 percent (see Figure 3, redrawn by the author from Gu–Kelly–Xiu 2020, Table 1). The monthly figures reported in the original paper are the following: PLS ≈ 0.27, PCR ≈ 0.26, ENet ≈ 0.11, RF ≈ 0.33, GBRT ≈ 0.34, NN3 ≈ 0.40.

Figure 3: Comparison of out-of-sample R^2 across the 13 predictive models (OLS+H omitted)



Redrawn by the author from Gu–Kelly–Xiu 2020, Table 1. Note: The R^2 of the all-variable OLS model (OLS+H) is −3.46 percent; it is not plotted because of its excessively large negative value. The remaining 12 models are plotted in full.

Although the figure looks small, it should be read against the right benchmark: it is a predictive gain that is still substantial at the individual-stock level and on a monthly basis, and it can be converted into portfolios with fairly high out-of-sample Sharpe ratios; net returns should be evaluated separately from transaction costs.

A significant operational benefit of this format is that it bypasses the EIV problem of the Fama–MacBeth structure. Machine learning does not need a first-stage estimate of β and a second-stage estimate of λ; it predicts directly from features out of sample and therefore avoids the circularity of estimating one estimator from another. This differs from the EIV correction of Jegadeesh et al. [11] inside the classical framework: the latter improves the quality of the estimate within the original framework, whereas the former changes the evaluative question from whether pricing errors equal zero to whether out-of-sample returns can be predicted systematically.

The applicability of the out-of-sample evaluation framework has also been established in cross-asset settings. Chinco, Clark-Joseph, and Ye [13] use LASSO to identify sparse signals in the cross-section of individual stocks, and Bianchi, Büchner, and Tamoni [14] use machine learning to predict bond risk premia. In both cases, out-of-sample predictive power serves as a measure of validity outside equities.

The metric should be defined explicitly: the out-of-sample R^2 of about 0.4 percent is computed from monthly U.S. equity returns over 1957–2016 and compared across 13 predictive models. It is a measure of predictive performance and cannot be compared directly with pricing-error measures such as alpha or GRS; Sharpe ratios belong to a different class of portfolio-performance measures and should not be conflated with it. This confusion of metrics is common in the literature.

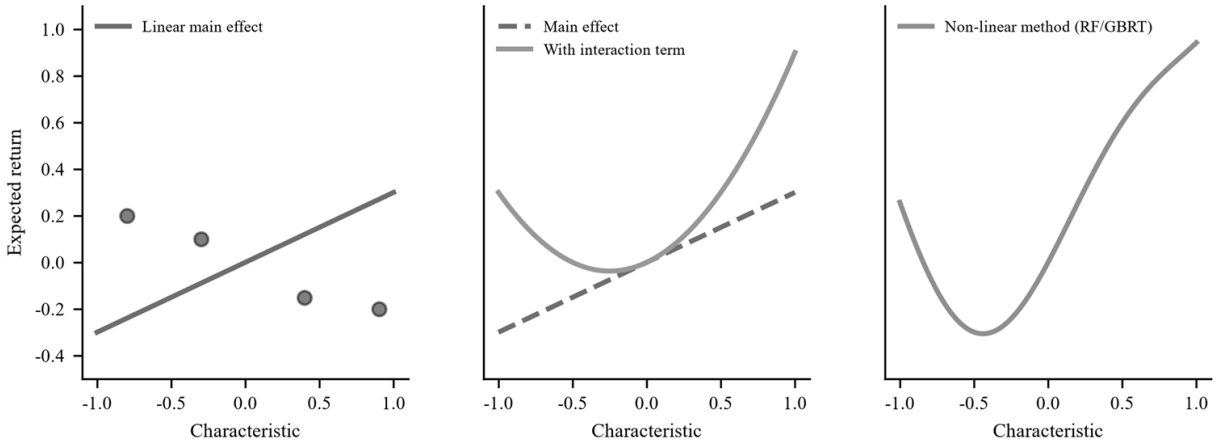
© 2026 The Author(s). Authors retain copyright in their articles.

3.3 Core Verification: Where Do the Gains of Nonlinear Methods Come From?

A deeper question needs to be answered: in which specific area do the benefits of machine learning come into play? The answer is not a matter of more complicated models but of two origins. Compared with linear parametric models (OLS, LASSO), nonparametric approaches can automatically learn the interactions between features and the nonlinearities in the data; compared with sorted portfolios, they reduce the loss of information caused by discretisation and by the impossibility of exhaustively enumerating high-dimensional interactions.

One important point is that if true returns are generated by interactive relationships among firm characteristics, linear parametric models (OLS, LASSO) represent main effects only and therefore carry no interaction information, whereas nonparametric techniques (random forests, gradient boosting) can generate interaction terms and nonlinear transformations automatically and thus deliver conditional gains (see Figure 4, a conceptual schematic).

Figure 4: Illustrative comparison between linear main effects and automatic interaction capture



Note: This is a conceptual diagram, not empirical data; it illustrates the interaction information that is lost when only main effects are modelled.

This logic matters because it turns the vague question of whether machine learning is more effective than linearity into a testable hypothesis: when features interact or the return structure is nonlinear, linear parametric methods can fail to capture information whereas nonparametric methods can achieve conditional gains. It is not a question of the superiority of algorithms but of what the evaluation model is designed to capture.

The limits should also be defined: the small improvement of nonlinear approaches holds only when the return structure really does contain interactions or nonlinearities. When the true data-generating process is linear, the benefit of nonlinear methods vanishes and can even turn negative because of overfitting. This limit leads directly to the criticism of overfitting in Direction VI and shows that the statement that "nonlinear is better than linear" is an empirical assessment, not a final verdict.

3.4 Section Summary

The Direction II argument may be summarised in three points. First, the challenges facing the traditional tests are structural: the EIV problem of Fama–MacBeth and the correlated nature of the test assets make the pass criterion excessively easy, a flaw that cannot be repaired merely by raising the t-value thresholds. Second, the machine-learning framework of out-of-sample evaluation provides a different standard for assessing factors; it avoids the EIV problem of the Fama–MacBeth framework, but out-of-sample R^2 reflects predictive ability and cannot be compared with pricing-error indicators such as alpha or GRS. Third, the extra benefit of nonlinear techniques is situational: it arises from the information loss of linear parametric techniques and from the operational constraints of sorted portfolios, and it applies only when interactions or nonlinearities are present in the return structure.

These three points support the thesis that the real value of machine learning is not confined to the release of a stronger prediction engine; machine learning offers an additional, practical standard for the notion of factor

© 2026 The Author(s). Authors retain copyright in their articles.

validity. Because in-sample significance is extremely vulnerable to scale pressure, introducing an out-of-sample evaluation framework is methodologically important; it remains, however, an addition to the existing concept rather than a replacement of it.

4. Changes in the Construction Paradigm: From Sorting to Automatic Interaction Discovery

The testing phase answers the question of whether a factor is valid; the construction phase answers the question of how it is produced. The main argument of this section is that the manual sorting procedures underlying conventional factor construction run into practical constraints in high-dimensional settings — discretisation loss, arbitrary cut points, the curse of dimensionality, and the difficulty of enumerating interactions exhaustively; machine learning addresses them by searching automatically for interactions between features and for nonlinear structures, but it lacks interpretability.

4.1 Limitations of Traditional Sorting Methods

Single- and double-sorted portfolios are the standard operations by which traditional factors are constructed: stocks are sorted on a characteristic, grouped, and the long–short return of the extreme groups is used as the factor return.

A popular misconception should be clarified beforehand: sorting is a nonparametric grouping technique and does not imply that the relationship between characteristic and return must be linear or monotonic. Decile sorts may be U-shaped, inverted U-shaped, or otherwise non-monotonic; double sorts consider return differences conditional on two dimensions and do not necessarily assume the absence of interaction between characteristics.

The genuine limitations of sorting methods are the following four.

First, discretisation causes information loss. Placing continuous characteristics into bins removes the within-bin variance and also lowers the precision of the factor returns.

Second, grouping breakpoints are subjective. The quintiles, deciles or any other divisions and their breakpoints are not standardised; inferences can vary depending on the grouping scheme.

Third is the curse of dimensionality. The number of cells in a multi-way sort grows exponentially with the number of characteristics, the sample size in each cell falls quickly, the precision of the estimates declines, and multi-characteristic interactions cannot be studied efficiently.

Fourth, it is hard to enumerate all possible multi-characteristic interactions. Manual sorting can cross two or three characteristics at best; interactions of a higher order are not feasible in practice.

The interaction effects among predictors are a significant contributor to the gains of tree models and neural networks, as Gu, Kelly, and Xiu [12] state. This is not equivalent to saying that traditional portfolio sorts assume linearity; it points rather to the operational limitation of sorting, namely its inability to represent higher-order interactions.

4.2 From Sorting to Dimension Reduction: Evolution within the Linear Path

One point that is easily misunderstood needs to be clarified. The operational constraints of sorting were discussed in Section 4.1 (information loss due to discretisation, arbitrary breakpoints, the curse of dimensionality, and the difficulty of enumerating multi-characteristic interactions).

Sorting is not, however, the only way of constructing factors. Econometricians had already suggested an alternative long before machine learning entered the subject, namely dimension reduction and extraction, that is, the direct derivation of latent factors from a characteristic panel without grouping. The three-pass regression filter (3PRF) of Kelly and Pruitt [15] is in fact a supervised dimension-reduction approach to forecasting with many predictors and is not, strictly speaking, an asset-pricing factor-construction technique; it is the IPCA of Kelly, Pruitt, and Su [16] that truly brought dimension reduction into factor construction.

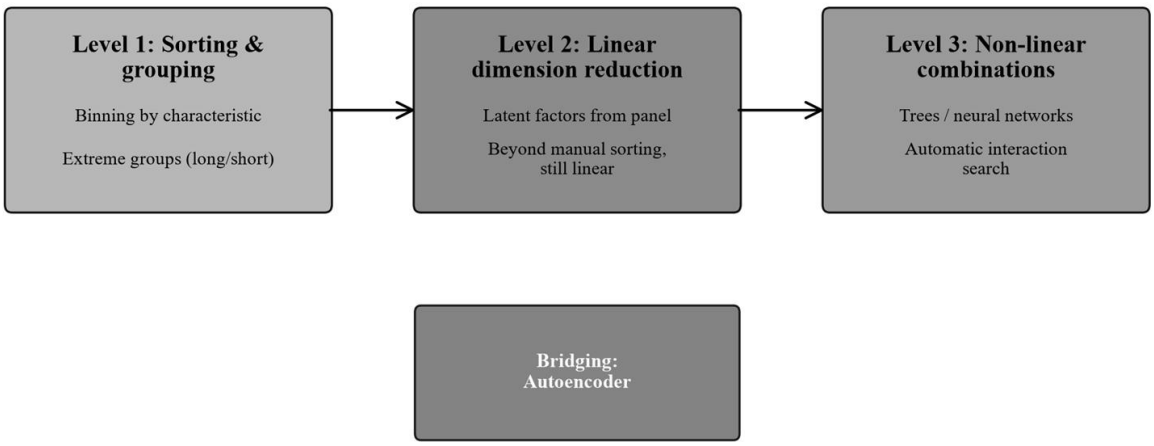
The most important difference is that dimension reduction addresses the discretisation loss of sorting but still resides in the linear-combination paradigm. It represents factors as linear functions of characteristics and

© 2026 The Author(s). Authors retain copyright in their articles.

cannot represent nonlinear relationships. Kelly, Pruitt, and Su also develop the idea that characteristics are covariances by transforming characteristics into time-varying factor loadings [16]. It should be observed, nevertheless, that the statement that characteristics are covariances is itself a hypothesis requiring further verification; it has not yet gained enough validation to be treated as universally applicable.

Between linear dimension reduction and nonlinear combination, the autoencoder asset-pricing model of Gu, Kelly, and Xiu [17] acts as a transitional bridge: the autoencoder compresses features and identifies latent factors in a nonlinear manner, thereby breaking the linearity presumption at the dimension-reduction step (see Figure 5, drawn by the author).

Figure 5: Evolution of the Factor-Construction Paradigm: From Linear Dimension Reduction to Nonlinear Combination via Autoencoders



Drawn by the author. Note: This is a conceptual diagram without data; it shows the three-tier evolution of the factor-construction paradigm.

4.3 Decision Trees and Automatic Interaction Capture

Tree-based methods such as random forests and gradient boosting are instances of automatic interaction discovery, since they do not rely on a predetermined set of interaction patterns but explore interactions between features and nonlinear partitions in a recursive manner.

The conclusion of Gu, Kelly, and Xiu [12] is that tree models and neural networks are the highest-performing of the 13 predictive models, which is exactly what this ability implies. Tree models can construct interactive features automatically, such as high turnover combined with low volatility, whereas manual sorting requires the researcher to see the interaction in advance.

Bryzgalova, Pelger, and Zhu [18], in “Forest Through the Trees”, go further and construct test assets with decision trees: instead of sorting manually on a size–B/M grid, the sorting itself is performed by decision trees. The importance of this work is that it extends the decision tree from a prediction tool to a test-asset construction method. It should be noted that the paper is a working paper that has not completed peer review; it counts as frontier exploration rather than as a settled conclusion.

This advance prompts a deeper question: since decision trees can uncover interactions that manual sorting misses, have conventional factors already lost information at the construction stage? At present it can be stated that when a return structure contains interactions and nonlinearities, traditional sorting is prone to information loss. This is an internal link between Direction III and Direction II: the problems of traditional tests (Direction II) may be rooted partly in the shortcomings of the construction approach with respect to discretisation loss and the enumeration of high-dimensional interactions (Direction III) — although such an evaluation is not obligatory.

© 2026 The Author(s). Authors retain copyright in their articles.

4.4 The Cost of Interpretability

The cost of this change in construction method concentrates on a single issue: interpretability.

Sorting and linear dimension reduction have an unassailable benefit: they can answer precisely the question of which characteristic, at what weight, drives returns. This makes the factor economically narratable, explainable by theory, intelligible to regulators, and trustworthy to practitioners.

Decision trees and neural networks are nonlinear and give up exactly this benefit. It is hard to compress the split paths of a trained random forest into a succinct economic mechanism, and the internal representations of a deep neural network cannot be converted into claims of the form "characteristic X determines return Y". This is a significant loss for a field whose central goal is mechanistic understanding.

The cost is not fixed. Feature-importance measures and SHAP (Shapley additive explanations) are interpretability methods developed in recent years that reduce the black-box problem to some degree, but they supply information about which feature is statistically important rather than why the feature influences returns in economic terms. The first is statistical explanation, the second is economic mechanism, and the two should not be confused. This distinction leads directly to the notion of economic retrospective validation developed in Direction IV.

4.5 Section Summary

The Direction III argument may be summarised in three points. First, traditional sorting does not rest on a linearity assumption, but discretisation loss, arbitrary breakpoints, the curse of dimensionality, and the difficulty of enumerating interactions all cause information loss. Second, the machine-learning construction process has two stages: the first is linear dimension reduction, which removes the discretisation loss of sorting, and the second breaks the linear-combination structure itself through nonlinear solutions. Third, the price of this transition is interpretability; nonlinear construction comes at the cost of the economic story conveyed by the characteristic–mechanism link, and statistical explanation is no substitute for economic mechanism.

These three points together formulate the question that Direction IV should answer: when factors are constructed by algorithms rather than by human hands, how does the source of knowledge about factors change? How can algorithmically extracted factors be shown to carry economic significance rather than being mere statistical artefacts?

5. The Discovery Path: Toward Data-Driven Methods and the Backward Shift of Theory's Role

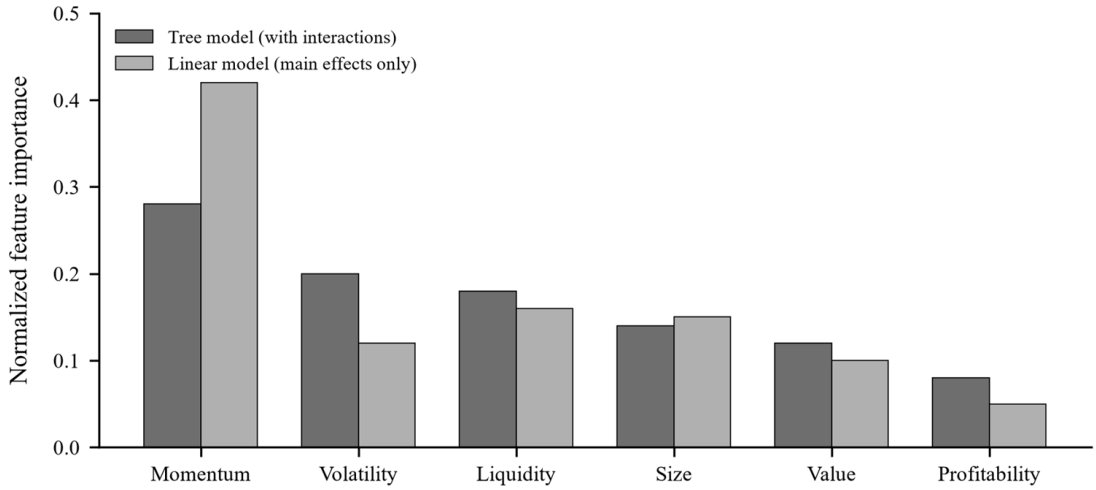
The construction phase alters the way in which factors are produced; the discovery phase alters the origin of factors. The main point of this section is that machine learning provides a data-based route of discovery, which leads to a partial shift of the source of knowledge from human prior theory to algorithmic search. This transition does not and will not remove the need for economic validation; rather, it changes the role of economics from that of a pre-existing hypothesis to that of a retroactive examination.

5.1 The Transfer of the Source of Knowledge

Conventional factor discovery follows one route: theory first builds the factor, which is then tested; knowledge starts from theory and data merely decide whether the theory passes. Machine learning offers a different route: signals are mined by data-driven methods in a high-dimensional feature library without the need to predict in advance which features are correct. The role of data in the production of knowledge therefore increases, and theory at certain points becomes a post hoc explanation (see Figure 6, a schematic by the author, not an empirical finding).

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 6: Illustrative comparison of feature-importance rankings between tree-based and linear models



Note: Illustrative data only. The tree model attributes larger contributions to more features (such as volatility and liquidity) than the linear model does.

The change is clearly visible in the methodological spectrum.

Variable selection is the first layer. The sparse selection introduced by Tibshirani [19] for high-dimensional features through penalised regression is one of the oldest forms of data-driven discovery. Chinco et al. [13] apply it to the cross-section of individual stocks, deriving sparse predictors from a large number of lagged-return signals.

The second layer is factor extraction. The IPCA of Kelly, Pruitt, and Su [16] does not require the economic meaning of the latent factors to be specified in advance; it estimates simultaneously the latent factors and their characteristic-driven time-varying loadings (although the number K of latent factors still has to be set). The advance is to move from selecting existing features to estimating latent factors.

The third layer is semantic and alternative-data signal mining. Loughran and McDonald [20] build a specialised finance dictionary and show that general-purpose dictionaries perform poorly on financial text, which makes a domain-specific dictionary a basic instrument for quantifying textual signals. Jiang, Kelly, and Xiu [21] extract trend signals from price-chart images with convolutional neural networks, extending the carriers of signals from numeric features to raw image data.

The next level is semantic knowledge from large language models. Lopez-Lira and Tang [22] use GPT-4 to read news headlines and anticipate the stock-price responses to them; the paper has been published in the Journal of Financial Economics.

A caveat on evidence: the most recent layers of this lineage (images, large language models) are largely new work, and some are still working papers or forthcoming. They are the leading directions of the discovery path, not its final ones. Citations should therefore be explicit about distinguishing evidence that has been published and peer-reviewed from evidence that is still a working paper or forthcoming, so that frontier exploration is not confused with mature discovery.

5.2 The Core Risk of the Discovery Path: Statistical Artefacts Masquerading as Economic Factors

Data-driven discovery has a fundamental risk: when the feature space is large enough, the algorithms may be able to discover some apparently meaningful factor of in-sample historical returns, but the factor might just be an artefact of noise.

This risk is not hypothetical; it can be observed. A plausible explanation for why 65 percent of the analysed anomalies could not be replicated by Hou, Xue, and Zhang [3] is that many of them were statistical artefacts

© 2026 The Author(s). Authors retain copyright in their articles.

created by high-dimensional searching rather than real factors explicable by economic mechanisms. As the discovery process shifts its emphasis from theory-first to data-first, this risk increases systematically.

Data-driven discovery should therefore include two non-negotiable steps: out-of-sample validation and economic-mechanism retrospection. The first examines the strength of the signal on a new data set (repeating the out-of-sample R^2 criterion of Direction II); the second explains the existence of the signal in terms of economic reasoning. Neither can be neglected: out-of-sample validation without mechanistic explanation is a black box that predicts without understanding, while mechanistic explanation without out-of-sample validation produces a story that is coherent but unreliable.

5.3 The Necessity of Economic Retrospective Validation

The threat identified in the previous subsection leads directly to the argument of this one: economics has not been pushed aside by the discovery path; it has moved to a different point in it.

In the conventional paradigm, economics stands at the beginning of the discovery chain: theory first puts forward hypotheses and data subsequently test them. On the route outlined by machine learning, economics brings up the rear: algorithms first extract signals and theory explains them afterwards. This backward shift may appear to undermine the power of theory, but in reality it enhances its importance. The number of signals extracted by algorithms is far larger than the number of hypotheses presented by people; without an economic mechanism to act as a filter, researchers would be overwhelmed by unintelligible statistical signals.

As Giglio, Kelly, and Xiu [23] state clearly, machine learning is not a panacea and does not replace economic theory. The judgment reached here is that machine learning has changed the effectiveness and the scope of what is called discovery, but not the academic baseline of the field, according to which a factor must be explicable through an economic mechanism.

5.4 Section Summary

The claim of Direction IV can be summarised in three points. First, the machine-learning discovery path partially displaces the source of knowledge from prior theory towards data-driven search, and its methodological span runs from LASSO variable selection and IPCA factor extraction to image- and LLM-based semantic mining. Second, the central danger of this turn is the appearance of statistical artefacts posing as economic variables: in a highly dimensional feature space, algorithms may end up fitting false signals that are non-trivially significant in sample. Third, the requirement of economic validity has not been eliminated but has moved into the category of retrospective testing — algorithms widen the range of discovery, theory screens and explains, and the two complement each other rather than competing.

6. Cross-Asset Extension of the Machine-Learning Framework

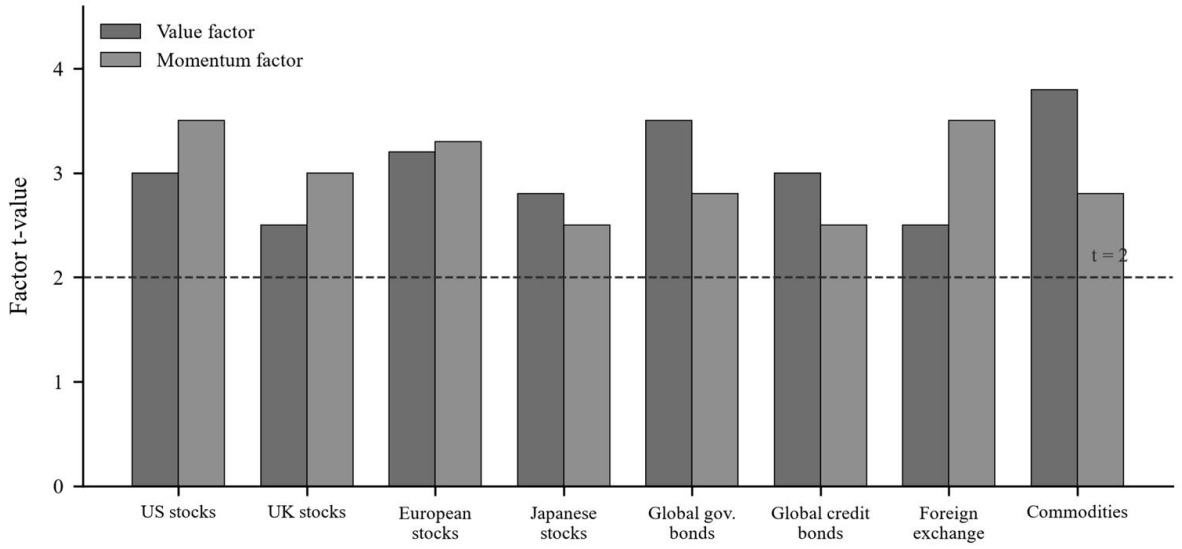
Directions II, III, and IV concern changes at the testing, construction, and discovery stages respectively, but they share one weakness of premise: these changes mostly take place in the equity asset class. Were the machine-learning paradigm to apply only to the U.S. equity cross-section, it would be of local importance. This section argues that if machine learning is to be applied to cross-asset domains, three further issues must be confronted: whether factor isomorphism holds across assets, whether the methodological challenges of joint estimation can be overcome, and how institutional differences limit the transfer of models. These points are emphasised in order to draw a sharp line between the contribution of machine learning and the existing cross-asset literature, so that the findings of Asness et al. [24] on cross-asset factors, those of Giglio and Xiu [27] on omitted factors, and the Chinese CH-3 model [28] are not attributed directly to machine learning. The discussion of the Chinese market serves as an illustration of institutional dependence and transfer boundaries, not as proof that a unified framework has been created.

6.1 Cross-Asset Factor Isomorphism: Value and Momentum Everywhere

The empirical basis of cross-asset factors is the work of Asness, Moskowitz, and Pedersen [24]. The article examines value and momentum factors in eight different markets and asset classes and shows that they are both prevalent and positively correlated across markets and asset classes (see Figure 7, redrawn by the author from Asness et al. 2013).

© 2026 The Author(s). Authors retain copyright in their articles.

Figure 7: The prevalence of value and momentum across eight markets and asset classes



Redrawn by the author from Asness–Moskowitz–Pedersen 2013 data. Note: The two factors are consistently significant and positively correlated across asset classes, which supports a unified framework.

The theoretical implication of this result is that if value and momentum are both present across markets and asset classes, they are not phenomena specific to a particular asset but cross-asset pricing regularities. This provides an empirically credible justification for a unified factor framework in which factors are no longer seen as unique to equities but as common pricing factors across asset classes.

The same isomorphism is confirmed in fixed-income assets. Cochrane and Piazzesi [25] show that a single return-forecasting factor constructed from forward rates predicts bond excess returns; Chordia, Goyal, Nozawa, and Subrahmanyam [26] show that several equity anomalies are also present in corporate bond returns. Taken together, these works suggest that cross-asset factor isomorphism may be a general phenomenon in asset pricing, although the magnitude of a particular factor differs across assets, and citations should be careful about the strength of the claim.

As a caveat on evidence, cross-asset isomorphism has been documented in research covering different asset types, frequencies, and sample periods. These studies agree on the direction of the finding (factors are present across asset classes), but the claim that value holds in commodities and the claim that it holds in equities cannot be treated as resting on the same weight of evidence.

6.2 Unified Estimation under Omitted Factors: The Theoretical Basis of Integration

Cross-asset isomorphism implies that a common framework can be developed, but it addresses only whether such a framework is feasible, not how it should be implemented. The conceptual barrier to a unified framework is that when a researcher tries to estimate a unified stochastic discount factor (SDF), any omitted pricing factor biases the estimate. This is exactly the issue tackled by Giglio and Xiu [27].

As shown by the three-step procedure of Giglio and Xiu [27] in "Asset Pricing with Omitted Factors", the risk prices of observable factors can be estimated consistently even in the presence of omitted factors. The importance of the work lies in providing a methodological basis for estimating a common SDF: even when the researcher cannot take all factors into account, consistent estimates remain achievable in the presence of omitted ones.

It should be made clear that there is a connection between Directions II and III: the omitted-factor model of Giglio–Xiu, the machine-learning-based out-of-sample performance evaluation of Direction II, and the latent-factor model of IPCA in Direction III offer possible methodologies for the three questions underlying cross-asset unification — what is omitted, how to predict performance, and how to build factors. These are three interrelated steps; they do not imply that machine learning triggered cross-asset unification.

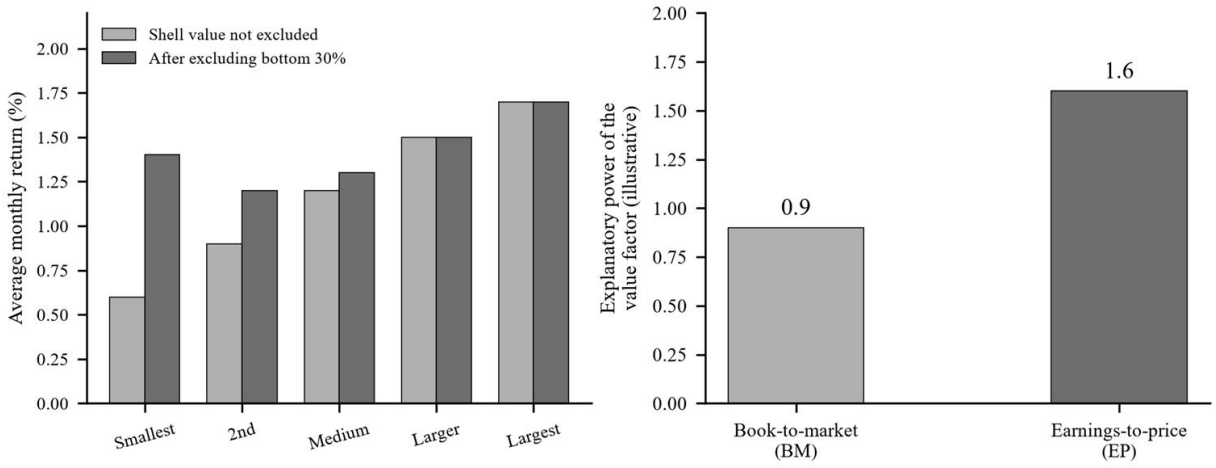
© 2026 The Author(s). Authors retain copyright in their articles.

6.3 The Chinese Market as an Illustration of Institutional Dependence and Transfer Boundaries

The final test of a unified structure is not whether it fits U.S. equities but whether it applies to other markets with very different institutional environments. From this point of view, the Chinese market provides important evidence: institutional differences have non-trivial effects on the construction and the performance of factors.

The Chinese three-factor model (CH-3) proposed by Liu, Stambaugh, and Yuan [28] illustrates the essence of this specificity: the smallest 30 percent of A-share market-capitalisation stocks are contaminated by shell value, which means that their prices are substantially affected by the reverse-merger option value rather than by operating fundamentals. Including this group would therefore create a systematic bias in the size effect. At the same time, book value under Chinese accounting rules is quite noisy, so the earnings-to-price ratio (EP) should be used instead of book-to-market (BM) as the value measure (see the schematic in Figure 8, reproduced by the author from the adjustment scheme of Liu–Stambaugh–Yuan 2019; it is not an empirical regression result).

Figure 8: Chinese-market institutional calibration: CH-3 adjustments relative to FF



Schematic drawn by the author from the adjustment scheme of Liu–Stambaugh–Yuan 2019. Note: The chart is schematic. The key insight is that excluding shell value and using EP instead of BM are essential steps of institutional calibration.

Excluding the lowest 30 percent by market capitalisation and replacing BM with EP are not merely data-manipulation techniques but essential adjustments through which institutional specificity enters the pricing model. It follows that a machine-learning framework cannot be transplanted into the Chinese market directly from the U.S. equity model; institutional calibration must come first.

Empirical work published in Chinese supports this assessment. Li et al. [29] find that the Fama–French five-factor model outperforms the three-factor model in the Chinese equity market, and that the profitability and investment factors become priced after the split-share structure reform; Zhao et al. [30] provide a systematic comparison of the five-factor and three-factor models. All of this work implies that a unified approach can be implemented in the Chinese market only through institutional calibration, and that institutional calibration itself shows cross-market transfer to be a conditional rather than an unconditional generalisation.

This judgment is a contribution of the paper: to our knowledge, English-language reviews rarely examine the Chinese-market evidence, which shows that transferring the framework first requires institutional adjustment. It is at once a test of the unified framework and a delineation of the boundary beyond which it can no longer be claimed to generalise without qualification.

6.4 Section Summary

Direction V can be summarised in three statements. First, cross-asset factor isomorphism provides empirical support for a unified framework but cannot be equated with machine learning driving cross-asset

© 2026 The Author(s). Authors retain copyright in their articles.

unification. Second, methods of unified estimation in the presence of omitted factors provide a theoretical basis for constructing a unified SDF; they relate to machine learning as candidate tools for the same problem domain rather than as extensions of machine learning results. Third, because of its institutional particularity, the Chinese market provides significant information about the limits of transferring the machine-learning framework.

7. Evaluating the Evidence: Are the Gains of Machine Learning Worth Their Costs?

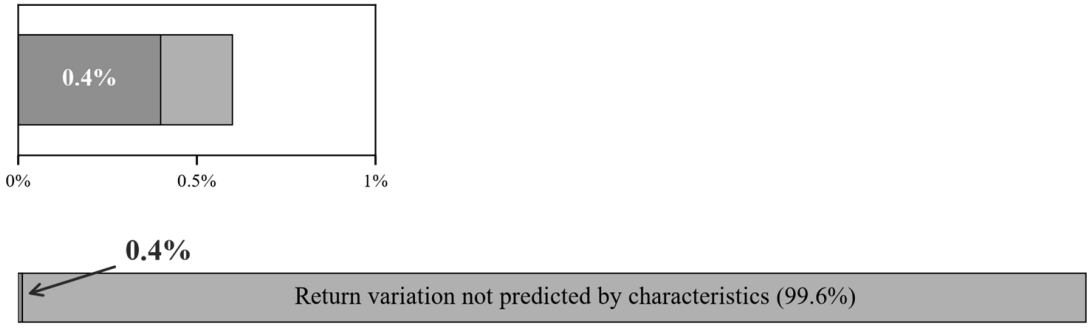
It would be easy to slip into an unexamined, optimistic conclusion from the arguments of Directions I through V: that machine learning is now comprehensively and unconditionally better than the old paradigm. The aim of this section is to stop that slide. The main assertion is that the advantages of machine learning are genuine but restricted, and that it carries losses (interpretability loss, conflation of correlation with causation, and overfitting) that are substantive rather than purely technical; machine learning therefore does not constitute an unconditional endorsement of the approach but a conditional judgment.

7.1 The Limited Nature of the Gains: The Meaning of an Out-of-Sample R^2 of 0.4 Percent

The first point to acknowledge openly is that the total gains of machine learning in predicting cross-sectional returns are limited. Gu, Kelly, and Xiu [12] report an out-of-sample R^2 of about 0.4 percent per month. This number was cited in Direction II to show that an out-of-sample evaluation model can be implemented; here it must be revisited from the opposite perspective: even the most effective machine-learning approaches provide only a relatively low degree of out-of-sample predictive capability. It is important to note that the out-of-sample R^2 of Gu et al. [12] takes the sum of squared un-demeaned excess returns as the denominator and a zero forecast as the benchmark; it does not carry the same meaning as the proportion of variance explained in an ordinary regression and therefore cannot be interpreted directly as explaining 0.4 percent of return variation.

This statement must not be misread either as saying that machine learning is ineffective or as exaggerating its value. It has two interpretations. First, although an out-of-sample R^2 of 0.4 percent is a very small number, the resulting predictive ranking may still generate portfolios with fairly high out-of-sample Sharpe ratios; machine-learning strategies, however, tend to have high turnover and should therefore be evaluated on net returns after transaction costs. Second, the absolute value implies that machine learning remains far from solving the central challenge of asset pricing: most of the cross-sectional variation in returns cannot be predicted from characteristics (see Figure 9, a conceptual schematic).

Figure 9: The "limited nature" of the 0.4% out-of-sample R^2



Conceptual schematic. Note: The 0.4 percent figure is an out-of-sample R^2 calculated relative to a zero-prediction benchmark, using the sum of squared un-demeaned excess returns as the denominator. It does not correspond to the "proportion of variance explained" of standard regression analysis. The bar length in this chart is for visual illustration only and should not be read as "machine learning explains 0.4 percent of return variance". This 0.4 percent is real yet limited: it is enough to form a tradable strategy, but it falls far short of solving the core challenge of asset pricing.

This double meaning is the starting point of Direction VI: if the gain is only 0.4 percent, the claim that machine learning has changed empirical asset pricing must be restricted to the precise statement that it has altered approaches without resolving the underlying problem.

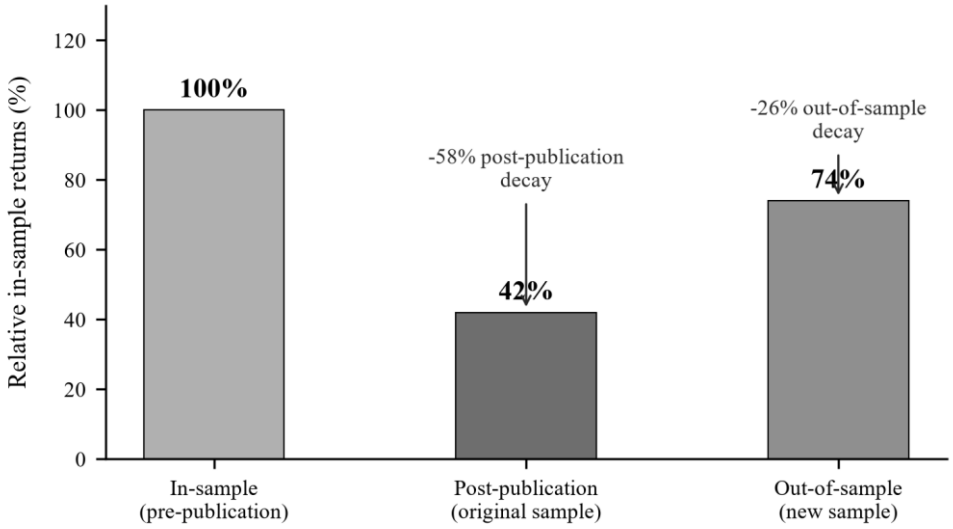
© 2026 The Author(s). Authors retain copyright in their articles.

7.2 Decay Risk: Out-of-Sample Predictive Power Is Not Naturally Immune to Publication Bias

One fact must be accepted: the out-of-sample predictive ability of machine learning is not naturally immune to publication bias and decay.

McLean and Pontiff [7] observe that post-publication returns on factors are 58 percent below in-sample returns. Although that result concerns conventional factors, the same process applies to signals discovered by machine learning: every published strategy or factor has passed through researcher selection, and such selection inevitably promotes the outcomes that worked best in sample (see Figure 10, reproduced by the author from McLean–Pontiff 2016).

Figure 10: Return decay after factor publication



Redrawn by the author from McLean–Pontiff 2016. Note: Post-publication decay is 58 percent and out-of-sample decay is 26 percent; the decay appears systematically after publication, suggesting that the publication process filters for the best statistical artefacts in the sample.

A deceptive myth surrounds machine learning on this point: out-of-sample validation is treated as a vaccine against overfitting. The truth is different. Out-of-sample validation can prevent overfitting to a fixed sample, but it does not protect against overfitting the validation step itself; when a researcher can keep changing the model, the features, and the validation windows until out-of-sample performance looks good, out-of-sample validation becomes a form of in-sample fitting.

The essence of the critique is that out-of-sample R^2 is a harsher condition than in-sample significance; it is not the end of overfitting, however, but a new front line that must be monitored continuously. The efficacy of the out-of-sample evaluation framework set out in Direction II will depend on the rigour of the validation protocol itself, not on any inherent protection of the standard.

7.3 Three Substantive Costs: Interpretability, Causality, and Overfitting

The restrictiveness of the gains and the risk of decay are quantitative concerns; machine learning also presents three qualitative costs that can affect the very principles of asset pricing as a field.

The first cost is interpretability. Section 4.4 already explained that nonlinear construction eliminates the economic story of the "characteristic–mechanism" form, which is the foundation of the discipline. Here it need only be connected with the other two costs: the loss of interpretability is substantive because it makes it harder to detect the conflation of correlation with causation, and overfitting itself.

The second cost is the confusion of correlation with causation. The out-of-sample predictive ability of machine learning answers the question of whether one can predict, not why one can predict. A black-box model with high predictive performance cannot be used to understand the causal process of pricing. The distinction

© 2026 The Author(s). Authors retain copyright in their articles.

is crucial because the ultimate objective of empirical asset pricing is not to forecast returns but to learn how risk is priced, which is a causal rather than a correlational hypothesis.

The third cost is overfitting. A sufficiently large feature space allows data-driven mining to find signals that are significant in sample yet fail to persist out-of-sample. This is not a problem that stricter validation can fix; it is a problem that accompanies the data-driven discovery path (in line with Direction IV).

Kelly, Palhares, and Pruitt [31], in "Modeling Corporate Bond Returns", present a possible response to these three costs: if machine learning is used together with economic constraints — that is, if the algorithm is embedded in an economic structure instead of being allowed to fit freely — predictability and interpretability can be achieved at the same time. The importance of this work is that it turns the contradiction between machine learning and economic theory into a complementarity, providing a methodology that leads to the conclusion of Direction VI.

7.4 Conditional Judgment: In What Sense Does the Contribution of Machine Learning Hold?

Synthesising Directions I through VI, this paper offers a conditional judgment.

Giglio, Kelly, and Xiu [23] state clearly in their survey that machine learning is not a panacea and not a replacement for economic theory. The precise interpretation of that judgment is the theoretical foundation of the conditional judgment offered here.

The conditional judgment can be stated as follows.

Machine learning has produced legitimate instrumental advances in prediction-oriented tasks: it has offered out-of-sample predictive performance as an additional standard for measuring factors, automatic interaction discovery as a way of building factors, and high-dimensional search as a realistic means of discovering factors. All three changes are real and structural.

With respect to causal inference and mechanistic understanding, however, machine learning has not been able to make a substitutive contribution: it can neither explain why risk is priced nor replace the explanatory role of economic theory, and its gains come at the cost of interpretability, causal reasoning, and robustness to overfitting.

It is therefore more accurate to say that machine learning has not transformed the empirical approach to asset pricing, but that it has changed the nature of the prediction-related processes in empirical asset pricing without altering the internal logic of the understanding-related processes. This distinction constitutes the conclusion of the paper. Table 1 summarises the correspondence between the traditional paradigm and the machine-learning paradigm on six task dimensions.

Table 1: Comparison of Tasks between the Traditional Paradigm and the Machine-Learning Paradigm

Task Dimension	Traditional Paradigm	Machine-Learning Paradigm	Substitution?
Validity judgment	In-sample significance (t-stat, Alpha)	Out-of-sample predictive performance (OOS R^2)	No
Factor construction	Manual sorted portfolios	Dimension reduction + automatic interaction discovery	Partial substitution
Factor discovery	Theory-prior hypotheses	High-dimensional signal search	Partial substitution
Pricing-error tests	Fama–MacBeth / GRS	No direct counterpart	No
Mechanism explanation	Economic-theory narrative	Feature importance / SHAP	No
Cross-asset unification	Factor isomorphism + SDF estimation	Can handle high dimensions, but faces transfer boundaries	No

7.5 Section Summary

The Direction VI argument can be summarised in three points. First, the achievements of machine learning are real but constrained: an out-of-sample R^2 of 0.4 percent is at once an indication of genuine effectiveness and an indication of how far the problem remains from being solved. Second, out-of-sample predictive capacity is not inherently immune to publication bias and decay, and out-of-sample validation can itself become a form of overfitting. Third, interpretability, causality, and overfitting are three substantive costs, and Kelly, Palhares,

© 2026 The Author(s). Authors retain copyright in their articles.

and Pruitt [31] show that these costs may be reduced by combining machine learning with economic restrictions.

8. Conclusion

By organising the argument along the four steps of testing–construction–discovery–integration, the paper has shown how machine learning has systematically changed the research tasks of empirical asset pricing and has reached a conditional conclusion.

The shifts in the four stages are real: the testing stage has established an out-of-sample evaluation benchmark (Direction II), the construction stage has opened a path of automatic interaction discovery (Direction III), the discovery stage has introduced data-driven search strategies (Direction IV), and the integration stage has led to the consideration of cross-asset unification and institutional dependence (Direction V). All these changes rest on the fact that the traditional paradigm cannot cope with scale pressure (Direction I): when up to 82 percent of anomalies cannot be reproduced under a strict protocol, the standard of "significant, therefore valid" is called into question.

The change nonetheless has limits (Direction VI). Although machine learning has made genuine instrumental progress on prediction-related problems, it has made no substitutive contribution to causal inference or mechanistic insight. The out-of-sample R^2 of 0.4 percent is an indication both that the gains are real and that the original problem persists; the three costs of interpretability, causality, and overfitting are substantive rather than merely technical.

The practical consequences of this conditional judgment are clear. For researchers, machine learning is useful not as a substitute for economic theory but as a tool that gradually turns theory from a prior hypothesis into a retrospective test: the algorithm widens the set of possible factors, and theory acts as a sieve and as explanation. Practitioners may use out-of-sample predictive performance as an operating standard, but should be careful about overfitting the out-of-sample validation step itself. In the Chinese market, implementing the machine-learning framework requires institutional calibration (filtering out shell value, substituting EP for BM) and cannot be a straightforward transplant of the U.S. equity model.

The precise meaning of the phrase "from factor zoo to factor refinery" is therefore not that algorithms have replaced theory, but that algorithms have transformed how knowledge is produced while theory remains the ultimate judge of its truth or falsity. The refinery refines prediction, not comprehension; and comprehension has always been the essence of asset pricing research.

References

[1] Cochrane, J. H. (2011). Presidential address: Discount rates. The Journal of Finance, 66(4), 1047–1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x.

[2] Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. The Review of Financial Studies, 29(1), 5–68. https://doi.org/10.1093/rfs/hhv059.

[3] Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. The Review of Financial Studies, 33(5), 2019–2133. https://doi.org/10.1093/rfs/hhy131.

[4] Chen, A. Y., & Zimmermann, T. (2022). Open source cross-sectional asset pricing. Critical Finance Review, 11(2), 207–264. https://doi.org/10.1561/104.00000112.

[5] Jensen, T. I., Kelly, B., & Pedersen, L. H. (2023). Is there a replication crisis in finance? The Journal of Finance, 78(5), 2465–2518. https://doi.org/10.1111/jofi.13249.

[6] Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. The Journal of Finance, 75(3), 1327–1370. https://doi.org/10.1111/jofi.12883.

[7] McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? The Journal of Finance, 71(1), 5–32. https://doi.org/10.1111/jofi.12365.

© 2026 The Author(s). Authors retain copyright in their articles.

[8] Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. Journal of Political Economy, 81(3), 607–636. https://doi.org/10.1086/260061.

[9] Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. Econometrica, 57(5), 1121–1152. https://doi.org/10.2307/1913625.

[10] Lewellen, J., Nagel, S., & Shanken, J. (2010). A skeptical appraisal of asset pricing tests. Journal of Financial Economics, 96(2), 175–194. https://doi.org/10.1016/j.jfineco.2009.09.001.

[11] Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. (2019). Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. Journal of Financial Economics, 133(2), 273–298. https://doi.org/10.1016/j.jfineco.2019.02.010.

[12] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009.

[13] Chinco, A., Clark-Joseph, A. D., & Ye, M. (2019). Sparse signals in the cross-section of returns. The Journal of Finance, 74(1), 449–492. https://doi.org/10.1111/jofi.12733.

[14] Bianchi, D., Büchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. The Review of Financial Studies, 34(2), 1046–1089. https://doi.org/10.1093/rfs/hhaa062.

[15] Kelly, B. T., & Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. Journal of Econometrics, 186(2), 294–316. https://doi.org/10.1016/j.jeconom.2015.02.011.

[16] Kelly, B. T., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics, 134(3), 501–524. https://doi.org/10.1016/j.jfineco.2019.05.001.

[17] Gu, S., Kelly, B., & Xiu, D. (2021). Autoencoder asset pricing models. Journal of Econometrics, 222(1), 429–450. https://doi.org/10.1016/j.jeconom.2020.07.009.

[18] Bryzgalova, S., Pelger, M., & Zhu, J. (2020). Forest through the trees: Building cross-sections of stock returns. Working paper. https://ssrn.com/abstract=3493458.

[19] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B, 58(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[20] Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x.

[21] Jiang, J., Kelly, B., & Xiu, D. (2023). (Re-)Imag(in)ing price trends. The Journal of Finance, 78(6), 3193–3249. https://doi.org/10.1111/jofi.13268.

[22] Lopez-Lira, A., & Tang, Y. (2026). Can ChatGPT forecast stock price movements? Return predictability and large language models. Journal of Financial Economics, 184, 104335. https://doi.org/10.1016/j.jfineco.2026.104335.

[23] Giglio, S., Kelly, B., & Xiu, D. (2022). Factor models, machine learning, and asset pricing. Annual Review of Financial Economics, 14, 337–368. https://doi.org/10.1146/annurev-financial-101521-104735.

[24] Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. The Journal of Finance, 68(3), 929–985. https://doi.org/10.1111/jofi.12021.

[25] Cochrane, J. H., & Piazzesi, M. (2005). Bond risk premia. American Economic Review, 95(1), 138–160. https://doi.org/10.1257/0002828053828581.

[26] Chordia, T., Goyal, A., Nozawa, Y., & Subrahmanyam, A. (2017). Are capital market anomalies common to equity and corporate bond markets? An empirical investigation. Journal of Financial and Quantitative Analysis, 52(4), 1301–1342. https://doi.org/10.1017/S0022109017000515.

[27] Giglio, S., & Xiu, D. (2021). Asset pricing with omitted factors. Journal of Political Economy, 129(7), 1947–1990. https://doi.org/10.1086/714090.

© 2026 The Author(s). Authors retain copyright in their articles.

[28] Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. Journal of Financial Economics, 134(1), 48–69. https://doi.org/10.1016/j.jfineco.2019.03.008.

[29] Li, Z., Yang, G., Feng, Y., & Jing, L. (2017). An empirical test of the Fama–French five-factor model in the Chinese stock market [in Chinese]. Journal of Financial Research, 444(6), 191–206.

[30] Zhao, S., Yan, H., & Zhang, K. (2016). Does the Fama–French five-factor model outperform the three-factor model? Evidence from the Chinese A-share market [in Chinese]. Nankai Economic Studies, (2), 41–59.

[31] Kelly, B. T., Palhares, D., & Pruitt, S. (2023). Modeling corporate bond returns. The Journal of Finance, 78(4), 1967–2008. https://doi.org/10.1111/jofi.13233.

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/).

© 2026 The Author(s). Authors retain copyright in their articles.

Financial Economics Research; Vol.3, No.4; 2026
E-ISSN: 3006-3663 P-ISSN: 3006-3655
DOI: https://doi.org/10.70267/fer.v3n4.162176
Published by: Zeus Press

Research on the Mechanism in the Transformation of BYD's Strategic Advantages into Financial Performance: Based on the Harvard Analysis Framework

Yuyue Zhang*

College of Economics and Management, Northeast Forestry University, Harbin, Heilongjiang 150040, China

**Corresponding author: Yuyue Zhang.*

Abstract

As competition in the new energy vehicle (NEV) industry continues to intensify, how an enterprise's strategic advantages are transformed into financial performance has become an important issue for the enterprise's sustainable development. Based on the Harvard Analysis Framework (HAF), this study takes BYD Company Limited (BYD) as the research object, carries out research from four aspects, namely strategic analysis, accounting analysis, financial analysis, and prospect analysis, and further discusses the transmission path through which corporate strategic factors affect financial performance. The study finds that corporate strategic advantages need to be transmitted through resource allocation, operating activities, and accounting matters, and are ultimately reflected in changes in financial indicators. On this basis, this study proposes suggestions such as improving the operating efficiency of the industrial chain, optimizing the transformation of research and development (R&D) achievements, and reasonably controlling the risks of international expansion.

Keywords

Harvard Analysis Framework, BYD, new energy vehicles, financial performance, strategic factors, transmission mechanism

1. Introduction

In recent years, the NEV industry has developed rapidly, the scale of China's NEV market has continued to expand, the industrial chain system has been continuously improved, and industry competition has gradually shifted from competition in market scale to comprehensive competition in technology, cost, and globalization capability. Against this background, whether an enterprise can effectively transform the strategic advantages it has formed into sustained and stable financial performance has become an important factor affecting the long-term development of NEV enterprises.

As an important enterprise in China's NEV industry, BYD has formed relatively strong competitive advantages in the NEV market by relying on its accumulation of power battery technology, industrial-chain integration capability, and continuous R&D investment [1]. In recent years, the company has strengthened supply-chain control through vertical integration, improved product competitiveness through technological innovation, and actively promoted its overseas market layout, achieving rapid improvement in corporate scale

© 2026 The Author(s). Authors retain copyright in their articles.

and market influence. However, as competition in the NEV industry intensifies, factors such as price competition, pressure from capacity expansion, and increased investment in overseas markets have gradually emerged. Whether the enterprise's strategic advantages can continue to be transformed into financial advantages is still worthy of further study.

Traditional corporate financial analysis usually takes financial indicators such as profitability, solvency, operating capability, and growth capability as the basis for evaluation. It can reflect the results of business operations, but it is relatively difficult to explain the strategic reasons behind the formation of financial results. The HAF breaks through the limitation of evaluating enterprise value by relying only on financial indicators, and combines strategic analysis, accounting analysis, financial analysis, and prospect analysis, providing a systematic method for understanding the relationship between business activities and financial performance. Therefore, further analysis based on the HAF of how corporate strategic factors affect financial indicators through operating activities and accounting matters can more comprehensively evaluate the development quality of NEV enterprises.

On this basis, this study takes BYD as the research object and, based on the HAF, conducts research from four aspects: strategic analysis, accounting analysis, financial analysis, and prospect analysis. It focuses on the influence paths of the industrial-chain integration strategy, technological innovation strategy, and international expansion strategy on financial performance. This study aims to analyze how corporate strategic advantages are transmitted to indicators such as profitability, growth capability, and financial structure through operating activities and accounting matters, and further proposes relevant suggestions for promoting sustainable development, so as to provide certain references for the financial analysis and strategic management of NEV enterprises.

2. Literature Review

The HAF breaks through the limitation of traditional financial analysis that focuses only on changes in financial indicators. It combines corporate strategy, accounting information, financial performance, and future development, providing a systematic analytical method for comprehensive corporate evaluation. In recent years, with the rapid development of the NEV industry, the HAF has gradually been applied to studies on the financial analysis of NEV enterprises. Li Yonghao and Hao Erhu analyzed BYD's financial performance by combining the HAF, providing a research foundation for the financial evaluation of NEV enterprises.

In terms of the formation of corporate competitive advantages, existing studies mainly focus on strategic choices and competitive capabilities of NEV enterprises. Chen Aizhen et al. pointed out that the development of the NEV industry depends not only on complete-vehicle manufacturing capability, but is also affected by industrial-chain coordination capability and the efficiency of upstream and downstream resource allocation. Qian Yisheng et al., based on a case study of BYD, found that vertical integration can improve the efficiency of internal resource allocation and strengthen the enterprise's control over key production links, providing support for the formation of long-term competitive advantages.

In terms of the relationship between technological innovation and corporate performance, existing studies focus on the impact of R&D investment on enterprise value and financial performance. Yang Zhonghuan found that R&D investment can affect an enterprise's value-creation capability, but its economic effect needs to be reflected through the transformation of innovation achievements. This indicates that the technological innovation strategy does not directly generate financial returns, but needs to go through the process of achievement transformation.

In addition, as NEV enterprises continue to promote global development, the influence of internationalization strategies on corporate financial performance has also received attention. Wang Chongfeng and Wang Xianguo analyzed the relationship between the internationalization strategy and financial performance of NEV enterprises from a dynamic perspective, and pointed out that internationalized layout can expand the market space of an enterprise, while overseas investment, production-base construction, and market development also increase capital investment and affect the enterprise's asset structure and financial risk.

In summary, existing studies provide a relatively sound theoretical basis for the strategic analysis and financial evaluation of NEV enterprises, but there are still certain deficiencies. Existing studies mainly

© 2026 The Author(s). Authors retain copyright in their articles.

concentrate on the analysis of corporate strategic advantages or the evaluation of financial indicators, and pay insufficient attention to the specific process by which strategic advantages further affect financial performance through resource allocation, operating activities, and accounting matters. Therefore, based on the HAF and taking BYD as the research object, this study further analyzes the transmission mechanisms through which the industrial-chain integration strategy, technological innovation strategy, and international expansion strategy affect financial indicators, so as to supplement the insufficiency of existing research on the linkage path between strategic factors and financial performance.

3. Strategic Analysis

3.1 Macro-environment Analysis of the NEV Industry: Political, Economic, Social, and Technological (PEST) Analysis

3.1.1 Political Environment

With the proposal of the dual-carbon goals, the NEV industry has gradually become a national strategic emerging industry and has helped realize the development of China's green economy. In recent years, China has continuously promoted the development of the NEV industry through a number of relevant policies supporting the industry. Since 2009, China has started subsidy actions for NEVs and issued the Adjustment and Revitalization Plan for the Automobile Industry. In 2013, the state issued the Notice on Continuing to Carry Out the Promotion and Application of New Energy Vehicles, and China's NEV industry began to develop vigorously. Among the six key tasks proposed in the Implementation Plan for Improving the System and Mechanism for Promoting Consumption (2018-2020), developing and strengthening green consumption and promoting the innovative development of intelligent vehicles were included. The vehicle purchase-tax exemption for NEVs and the NEV dual-credit policy also express the state's strong support for the new energy industry. However, with the gradual development of the industry, national subsidy policies have continuously raised standards and put forward higher requirements for product quality. The development of BYD's NEV business cannot be separated from the support of national policies. The company actively responds to national policies, strictly complies with corresponding laws, regulations, and policy norms, and formulates scientific and reasonable strategic decisions.

3.1.2 Economic Environment

China's economy maintains medium-high-speed growth, and its economic growth rate ranks first among the world's five largest economies. In 2025, gross domestic product (GDP) was significantly larger than in 2024, increasing by about RMB 5.38 trillion, and the annual real growth rate remained at 5%. Rapid economic development has made people pursue a higher-quality life. At the same time, China has actively promoted strategic adjustment of its economic structure, deeply advanced supply-side structural reform, and made green development deeply rooted in public awareness. In addition, the turbulent international situation and the outbreak of wars have caused gasoline and other energy prices to rise sharply, further increasing demand for NEVs. Moreover, the number of automobiles per thousand people in China has increased year by year, reaching 260 in 2025. However, compared with developed countries abroad, China's automobile ownership is still at a relatively low level, only one third of that in developed European and American countries. This shows that China's automobile industry still has large market space, and the future of NEVs is worth expecting. As a leading enterprise among China's NEV companies, BYD enjoys high-quality resources and a favorable economic environment.

3.1.3 Social Environment

With the continuous development of industry, protecting the natural environment has gradually become an important issue, and the concept of green travel has been deeply rooted in people's minds. Strong support for NEVs from government institutions and other bodies has greatly increased consumer acceptance of NEVs. According to data from the China Association of Automobile Manufacturers, in December 2025, China's NEV production and sales continued to grow year on year, and sales of new NEVs accounted for more than 50% of total new-vehicle sales. In addition, China's fuel demand mainly depends on imports. In recent years, wars have frequently broken out in some Middle Eastern countries, causing fuel prices to continue rising. It is

© 2026 The Author(s). Authors retain copyright in their articles.

necessary to further develop new energy, achieve self-sufficiency, reduce dependence on imported fuel, and improve national competitiveness. This provides BYD with a favorable social and cultural environment.

3.1.4 Technological Environment

China's scientific and technological level has developed rapidly in recent years. New technologies and new products have emerged one after another, and technological barriers have been continuously broken through, bringing many opportunities to the NEV market. The core of NEV R&D lies in motors and batteries. As a traditional battery manufacturer, BYD has a great advantage over its peers in battery R&D. The Blade Battery fully developed by BYD has successfully solved the range problem of NEVs. Strategies such as "524" formulated by the company have helped BYD achieve major technological innovation and improve its position in the industry [2]. In addition to battery technology, BYD has also continuously innovated in charging-pile technology and other areas, improving product competitiveness.

3.2 Market Analysis

The NEV industry has shown explosive growth in recent years. China, Europe, and the United States, as the main NEV markets, have always maintained the top three positions in global NEV market share. NEV sales reflect the development of the industry. In the past five years, global NEV sales have continued to rise. From the domestic perspective, China's NEV industry has grown rapidly, from 3.52 million vehicles in 2021 to 16.49 million vehicles in 2025. BYD and Tesla are close in sales and have also attracted much attention in recent years. The following part compares and analyzes the two companies. BYD was established earlier than Tesla, and its position gradually rose after its transformation toward NEV development. BYD designs diverse models and covers a relatively broad consumer group. It mainly targets middle- and low-end consumer groups and has gradually expanded toward high-consumption groups in recent years. Tesla faces high-income groups and benchmarks against high-end car companies in the international market. The gap between the two mainly lies in brand value, which is also where BYD is disadvantaged. Low brand value may cause BYD to be regarded as a middle- and low-end brand and not trusted when entering international markets such as Europe and the United States, and consumers may also have concerns about quality and prices in the used-car market. Therefore, BYD needs to improve product performance, provide good after-sales service, and optimize its path in the international market [3].

3.3 Industry Analysis: Porter's Five Forces Model

In terms of existing competitors, under the general environmental trend, automobile companies on the market and even some technology companies have gradually entered the NEV industry. Market competition is mainly reflected in price advantages and product differentiation. At present, BYD's competitors mainly include Tesla, Great Wall, NIO, and Geely. As a leading enterprise, BYD has always been the target pursued by other automobile companies. In 2023, the market concentration ratio (CR) of China's NEV industry was 39.51% for CR2, 52.63% for CR4, and 67.71% for CR8. Overall, the industry market concentration is relatively high, and the leading effect is obvious. In terms of potential entrants, as high-tech products, NEVs have relatively high requirements for technological R&D capability and expenses. Relevant national policies also create high entry barriers for this industry. In addition, for products such as automobiles that are related to daily personal safety, consumers are more inclined to trust well-known brands. BYD has operated for a long time and has already established a reliable image among consumers. It also has strong capital and strong technological R&D capability. Therefore, the threat of potential entrants is not significant. In terms of substitutes, the substitutes for NEVs are fuel vehicles. By comparison, fuel vehicles have a longer history and a more mature production system, and their production costs are relatively lower. However, with the development of the times, rising oil prices and increasing requirements for green environmental protection make reducing the use of fuel vehicles an inevitable trend. Although in some cold regions the endurance capability of NEVs is not as good as that of fuel vehicles, this must not be a barrier that technology cannot break through. BYD announced in March 2022 that it would completely stop producing fuel vehicles and focus on the NEV business. This decision is based on BYD's strong capability and profound understanding of the market. In terms of the bargaining power of buyers, BYD's main demand is divided into government demand and market demand. Government procurement aims to demonstrate and guide consumers to buy NEVs, and is not purely market behavior. Therefore, it is not sensitive to price and has weak bargaining power [4]. A large number of competitors and low-cost fuel vehicles have increased the bargaining power of ordinary buyers, but

© 2026 The Author(s). Authors retain copyright in their articles.

because of government subsidies and increasing public recognition of BYD, buyers' bargaining power will gradually decline. In terms of the bargaining power of suppliers, NEV components include the components required by traditional vehicles and batteries, motors, and other components. Components required by traditional vehicles, such as tires and glass, come from a mature industry with many suppliers, so suppliers' bargaining power is not strong. Battery and motor components are the core components of NEVs. For other enterprises, suppliers have strong bargaining power, but BYD started as a battery enterprise, owns multiple battery patents and mature battery technology, and is its own supplier.

3.4 Corporate Strategy Analysis

BYD is currently the only company in China with the technology for large-scale manufacturing of lithium iron phosphate vehicle batteries. Strong capital and self-sufficient raw materials are BYD's unique cost advantages. BYD's adoption of a cost-priority strategy can both form entry barriers for potential competitors and reduce buyers' bargaining power [5]. BYD adheres to a self-sufficiency route, insists on innovation, implements a diversified business strategy, avoids product homogenization with peers, and is determined to achieve full coverage in the automobile field.

3.5 Summary of Strategic Analysis and Its Financial Impact

The comprehensive analysis in this chapter shows that BYD's rapid development in recent years has benefited not only from the growth trend of the NEV industry, but is also closely related to the competitive strategy formed by the enterprise itself. Compared with traditional automobile enterprises, BYD has gradually formed relatively strong competitive advantages through the accumulation of power battery technology, industrial-chain layout, and continuous R&D investment.

According to resource-based theory, corporate competitive advantages come from the core resources and capabilities owned by the enterprise itself, but resource advantages need to realize value creation through business activities and organizational capabilities and ultimately appear as differences in corporate performance [6]. At the same time, the HAF emphasizes the relationship among the strategic environment, accounting information, and financial performance, providing a theoretical basis for analyzing an enterprise's financial status from a strategic perspective. Therefore, this study holds that BYD's industrial-chain advantages, technological advantages, and internationalized layout need to further reflect their financial impact through business activities and accounting matters.

On the one hand, the industrial-chain integration strategy [7] enhances the enterprise's control over core components and supply-chain links, helping reduce production costs and improve supply stability, thereby affecting profitability. However, as the enterprise expands in scale, industrial-chain construction also requires continuous investment in fixed assets and production resources, which may increase depreciation pressure and risks in asset utilization efficiency. Therefore, the subsequent accounting analysis needs to focus on changes in the scale of fixed assets, adjustment of asset structure, and the impact of related accounting estimates on financial performance.

On the other hand, technological innovation is an important basis for NEV enterprises to maintain competitive advantages. BYD has long invested in power batteries, intelligent technology, and core NEV technology, and the scale of R&D investment has continued to expand. R&D activities can enhance an enterprise's long-term competitive capability, but they also affect the enterprise's current profit level and involve accounting treatment issues such as capitalization and expensing of R&D expenditure. Therefore, the subsequent analysis needs to further focus on changes in R&D investment, the R&D capitalization ratio, and their impact on the enterprise's profitability and growth capability.

In addition, as competition in the NEV industry gradually shifts from a high-speed growth stage to a deepening competition stage, BYD has begun to actively promote overseas market layout and global development. International expansion can broaden the enterprise's future growth space, but the construction of overseas production bases, market development, and channel investment also increase the enterprise's demand for funds and may affect asset scale, debt structure, and solvency. Therefore, it is necessary to further analyze capital pressure and risk changes under the background of rapid expansion in combination with the enterprise's financial data.

© 2026 The Author(s). Authors retain copyright in their articles.

Overall, BYD's current financial performance is the result of the joint effects of the enterprise's strategic choices, changes in the industry environment, and business management. Therefore, in the following accounting analysis and financial analysis sections, this study further examines the economic effects and potential risks brought by BYD's strategic layout from the perspective of the linkage path between strategic factors and financial performance.

4. Accounting Analysis

4.1 Asset Analysis

Table 1: Asset Analysis Table

Year	2021	2022	2023	2024	2025
Total assets (RMB 100 million)	2958	4939	6795	7834	8837
Fixed assets (RMB 100 million)	1297	2531	3774	4128	5123

As shown in Table 1, from 2021 to 2025, BYD's asset scale continued to expand. From 2021 to 2023, the domestic NEV market was in an explosive stage. BYD built large numbers of complete-vehicle and power-battery production lines, planned to cancel complete-vehicle production of fuel vehicles, and fully produced NEVs. Total assets and fixed assets increased significantly. In 2024, they increased slightly, and in 2025, fixed assets increased again for overseas base expansion. The continuous growth of fixed assets over five years not only reflects BYD's strategic choice to expand capacity and promote industrial-chain layout, but also means that in the future the enterprise needs to absorb new asset investment through higher capacity utilization and market growth; otherwise, it may create pressure on asset turnover efficiency and profitability.

4.2 Liability Analysis

Table 2: Balance Sheet

Year	2021	2022	2023	2024	2025
Total liabilities (RMB 100 million)	1915	3725	5291	5847	6252
Short-term borrowings (RMB 100 million)	102	51.53	183.2	121	384.8
Long-term borrowings (RMB 100 million)	87.44	75.94	119.8	82.58	607.1
Total current liabilities (RMB 100 million)	1713	3333	4537	4960	4685

As shown in Table 2, BYD's total liabilities increased year by year. The growth of liabilities did not simply originate from operating pressure, but was related to strategic investments such as NEV capacity expansion, overseas layout, and industrial-chain construction. Therefore, analysis of liability changes needs to judge their rationality in combination with the enterprise's expansion stage. From 2022 to 2023, BYD's total liabilities nearly doubled, corresponding to large-scale construction of plants and other facilities during the explosive stage of the NEV market, and BYD supported expansion through debt financing. From 2023 to 2025, the rate of liability growth slowed and the pace tended to be stable. From 2021 to 2024, long-term borrowings remained low, and the company gave priority to short-term borrowings. In 2025, long-term borrowings increased significantly, short-term borrowings rose at the same time, and the enterprise started long-term capital expenditures such as overseas expansion.

4.3 Analysis of Key Accounting Policies and Accounting Estimates

BYD's main businesses involve NEVs, power batteries, automobile components, mobile-phone components, and other fields. They are characterized by a large scale of R&D investment, a relatively high proportion of fixed assets, and rapid product updating. Therefore, accounting policies such as R&D expenditure capitalization, fixed-asset depreciation, inventory impairment provision, and receivables impairment have a relatively large impact on understanding BYD's operating results and financial position [8].

First, from the perspective of R&D expenditure, the NEV industry has obvious technology-intensive characteristics, and enterprises need to continuously invest funds in the R&D of battery technology, complete-vehicle platforms, intelligent driving, and related core components. BYD adopts a technological innovation strategy and enhances its own advantages by continuously increasing R&D investment, while the accounting treatment of R&D expenditure directly affects the enterprise's current profit and asset scale [9].

© 2026 The Author(s). Authors retain copyright in their articles.

According to BYD's accounting policies, internal research and development projects are divided into the research stage and the development stage. Expenditure incurred in the research stage is directly included in current profit or loss. Expenditure in the development stage can be capitalized only when conditions such as technical feasibility, resource support, and reliable cost measurement are met.

In 2024, BYD's R&D expenditure was about RMB 54.161 billion, of which capitalized R&D expenditure was about RMB 966 million, with a capitalization ratio of about 1.78%. In 2025, R&D expenditure increased to about RMB 63.441 billion, capitalized R&D expenditure was about RMB 5.463 billion, and the capitalization ratio increased to about 8.61%. Although most R&D expenditure in 2025 was still expensed, the capitalized amount and capitalization ratio increased significantly, reflecting that some R&D projects gradually entered the achievement development stage.

The growth of R&D investment reflects BYD's strategic choice of technological competitive advantages, and attention also needs to be paid to the impact of R&D capitalization on financial information. On the one hand, R&D investment can promote the enterprise's long-term growth through technological breakthroughs, product upgrades, and improvement of market competitiveness. On the other hand, capitalized R&D expenditure reduces current expense recognition and increases current profit and asset scale. Therefore, when analyzing changes in BYD's profitability, attention needs to be paid to whether growth in R&D investment matches growth in profit, so as to avoid judging the enterprise's operating results only through profit indicators.

Second, from the perspective of fixed-asset depreciation, BYD adheres to the industrial-chain integration strategy and improves its independent control over core links through the layout of power batteries, component production, complete-vehicle manufacturing, and other links. At the same time, with the expansion of the NEV market, the company has continued to build production bases and promote overseas market layout, causing the scale of fixed assets to continue growing. Fixed-asset investment can enhance the enterprise's production capability and provide a basis for scaled operations, but it also brings problems such as increased depreciation costs and changes in asset utilization efficiency. BYD determines service life, estimated net residual value, and depreciation methods according to different asset categories, and regularly reviews related accounting estimates.

From 2021 to 2025, BYD's fixed-asset scale continued to increase, reflecting the enterprise's strategic investment in expanding NEV capacity, improving industrial-chain layout, and promoting global development. However, the NEV industry has a rapid pace of technological iteration, and some production equipment may face the risk of early elimination. If the estimated service life of fixed assets is higher than the actual economic life, depreciation expenses may be insufficiently recognized in earlier periods, causing the carrying amount of assets to be overstated. Therefore, when focusing on the competitive advantages brought by enterprise scale expansion, it is also necessary to further analyze whether new assets can be effectively transformed into capacity utilization and profitability.

From the perspective of inventory impairment provision, in 2025, BYD's inventory carrying balance was about RMB 142.622 billion, and the inventory impairment provision was about RMB 4.201 billion. The NEV industry has intense competition and rapid product updates. Changes in market prices may cause the value of some vehicle models, components, and raw materials to decline. Therefore, whether sufficient inventory impairment provision has been made directly affects asset quality and corporate profit.

Finally, expected credit loss of receivables is also an accounting policy that needs key attention. BYD has a large business scale and many customer types and business models. Risk judgment for accounts receivable and long-term receivables involve customer credit ratings, aging, historical collection, macroeconomic environment, and other factors. In 2025, the carrying amount of accounts receivable was about RMB 37 billion, and the carrying amount of long-term receivables was about RMB 24.4 billion. Although the enterprise has relatively strong collection ability by relying on its industry position and market competitive advantages, the difficulty of receivables management may also increase as sales scale expands and overseas business develops. If credit-risk assessment is insufficient, asset quality may decline and the stability of corporate cash flow may be affected [10]. Therefore, when analyzing BYD's growth capability, it is necessary to pay attention at the same time to the matching relationship between revenue growth and cash collection ability, so as to avoid pressure from fund occupation during enterprise scale expansion.

© 2026 The Author(s). Authors retain copyright in their articles.

4.4 Detection of Accounting Distortion and Accounting Adjustment

When analyzing corporate financial performance, attention needs to be paid to the possible impact of key accounting matters on financial information. In combination with BYD's technological innovation, capacity expansion, and market competition environment, this study focuses on matters such as R&D expenditure capitalization, inventory impairment provision, and credit risk of receivables.

First, from the perspective of R&D expenditure, BYD's R&D investment has continued to increase, and capitalized R&D expenditure in 2025 increased significantly compared with 2024. R&D capitalization conforms to the requirements of enterprise accounting standards, but changes in the capitalization ratio may affect the comparability of profits in different periods [11]. Therefore, when evaluating the enterprise's profitability, it is necessary to make a comprehensive judgment in combination with the scale of R&D investment and the transformation of R&D achievements, rather than evaluating the enterprise's operating results only on the basis of changes in current profit.

Second, from the perspective of inventory, intensified competition in the NEV industry and faster product updates may lead to a decline in the value of some vehicle models and components. BYD has made inventory impairment provision according to market conditions, but it still needs to pay attention to the impact of future market price changes on asset quality and profitability.

Finally, from the perspective of receivables, with the expansion of the enterprise's sales scale, the scale of accounts receivable increases accordingly. Although no obvious credit risk has been found at present, future overseas business expansion and business scale enlargement may increase pressure on fund management, and the recovery of accounts receivable needs continuous attention.

Overall, there is currently no obvious evidence of distortion in BYD's accounting information, and its financial statements can reasonably reflect the enterprise's operating conditions. However, factors such as changes in R&D capitalization, asset-scale expansion, and intensified market competition may still affect the results of financial analysis. Therefore, subsequent financial analysis needs to make a comprehensive evaluation in combination with strategic changes and the industry environment.

5. Financial Analysis

5.1 Profitability Analysis

Table 3: Profitability Analysis Table

Year	2021	2022	2023	2024	2025
Return on equity (ROE) (%)	3.73	16.14	24.4	26.05	15.31
Gross margin	13.02	17.04	18.58	19.44	17.74
Net profit margin	1.84	4.18	5.20	5.35	4.28

As shown in Table 3, from 2021 to 2024, BYD's return on equity (ROE), gross margin, and net profit margin all continued to increase, indicating that the ability of the enterprise's assets and capital to create profit was continuously strengthening. In 2025, all three profitability indicators declined. ROE fell to 15.31%, gross margin fell to 17.74%, and net profit margin fell to 4.28%. In combination with the preceding analysis, BYD adopted a vertical integration strategy to improve profitability by reducing supply-chain costs, but as price competition in the NEV industry intensified, the cost advantage was partly offset by price declines. Therefore, gross margin declined in 2025.

5.2 Solvency Analysis

Table 4: Solvency Analysis Table

Year	2021	2022	2023	2024	2025
Current ratio	0.970	0.722	0.666	0.747	0.793
Quick ratio	0.717	0.485	0.473	0.513	0.497
Asset-liability ratio (%)	64.76	75.42	77.86	74.64	70.74

The ideal values of a company's current ratio and quick ratio are 2 and 1 respectively. As shown in Table 4, BYD's current ratio continued to decline from 2021 to 2025 and was below 1 in all years, reflecting that the

© 2026 The Author(s). Authors retain copyright in their articles.

enterprise used more operating liabilities and financing to support scale growth during the high-speed expansion stage, and this needs to be judged comprehensively in combination with the enterprise's expansion strategy and cash-creation ability. The quick ratio decreased from 0.717 in 2021 to 0.497 in 2025, indicating that BYD's liquidity risk increased. The asset-liability ratio continued to rise from 2021 to 2023, indicating weak long-term solvency. From 2024 to 2025, the asset-liability ratio declined, and long-term solvency improved. Overall, BYD's solvency is relatively weak, but because BYD has relatively strong cash reserves, this will not affect its debt repayment [12].

5.3 Operating Capability Analysis

Table 5: Operating Capability Analysis Table

Year	2021	2022	2023	2024	2025
Inventory turnover ratio	4.244	4.869	5.360	6.026	5.146
Accounts receivable turnover ratio	5.58	11.30	11.96	12.52	16.19
Total asset turnover ratio	0.870	1.074	1.027	1.062	0.965

As shown in Table 5, from 2021 to 2025, BYD's inventory turnover ratio, accounts receivable turnover ratio, and total asset turnover ratio all increased overall, indicating improvement in the operating capability of assets and in asset utilization and liquidity. From 2021 to 2024, BYD's inventory turnover ratio continued to rise, reflecting strong demand in the NEV market. It declined in 2025, mainly because the enterprise increased stock preparation for overseas market layout and because the growth rate of industry demand slowed. The total asset turnover ratio continued to rise from 2021 to 2024 and declined in 2025 because of fixed-asset expansion. The accounts receivable turnover ratio increased significantly. Relying on BYD's leading position in the industry, the company has relatively strong bargaining power over dealers and others, greatly reducing bad-debt risk. From the vertical comparison, BYD's operating-asset management has greatly improved on its own basis [13]. Overall, BYD's operating efficiency improved compared with the beginning of the period from 2021 to 2025, but the inventory turnover ratio and total asset turnover ratio declined in 2025, indicating that the efficiency of its asset and inventory management still requires continuous attention under the background of expanded asset scale and inventory scale.

5.4 Growth Capability Analysis

Table 6: Growth Capability Analysis Table

Year	2021	2022	2023	2024	2025
YoY growth of net profit after deducting non-recurring gains and losses (%)	-57.53	1146.32	82.01	29.94	-20.38
YoY growth of net profit attributable to shareholders (%)	-28.08	445.86	80.72	34.00	-18.97
YoY growth of operating revenue (%)	38.02	96.20	42.04	29.02	3.46

As shown in Table 6, the year-on-year growth rates of operating revenue, net profit attributable to shareholders, and net profit after deducting non-recurring gains and losses can directly reflect the enterprise's scale expansion and profit-growth potential. From the perspective of operating revenue growth, BYD's growth shows obvious staged characteristics. From 2021 to 2022, due to the outbreak of the NEV industry, high-speed expansion was achieved, and the year-on-year growth rate of revenue reached 96.20% in 2022. Sales of complete-vehicle and power-battery businesses rose sharply, and market scale expanded rapidly. From 2023 to 2025, revenue growth continued to decline, and in 2025 it increased by only 3.46% year on year. The domestic NEV market entered the stage of stock competition, the industry price war intensified, overseas capacity layout was still in the investment cycle, and overall revenue expansion momentum weakened significantly. In 2021, rising raw-material prices and increased capacity-expansion costs caused corporate net profit to decline sharply year on year. In 2022, market demand was concentratedly released, economies of scale reduced production costs, net profit after deducting non-recurring gains and losses increased by 1146.32% year on year, and profitability achieved a leap-type recovery. From 2023 to 2024, profit growth gradually declined but still maintained positive growth. In 2025, affected by continuous price reductions in the industry and increased operating investment, the year-on-year growth rates of net profit attributable to shareholders and net profit after deducting non-recurring gains and losses both turned negative, and the enterprise's profit scale contracted.

© 2026 The Author(s). Authors retain copyright in their articles.

Taken together, the three growth indicators show that BYD fully grasped the industry trend in the early stage and achieved dual growth in scale and profit. However, in 2025, revenue growth nearly stagnated and profit declined year on year. The enterprise needs to rely on second growth curves such as overseas markets, energy storage, and photovoltaics to open growth space and alleviate the growth pressure brought by domestic stock-market competition. Whether overseas expansion can offset the decline in domestic market growth is an important factor in whether future strategic transformation can form new growth.

5.5 Transmission Mechanism Through Which Strategic Factors Affect Financial Indicators

Based on the HAF, the preceding sections analyzed and studied BYD's operating conditions from three aspects: strategic analysis, accounting analysis, and financial analysis. The analysis found that changes in BYD's financial performance are affected not only by the development environment of the NEV industry, but are also closely related to the enterprise's own strategic choices. Corporate strategic advantages are not directly manifested as financial results, but need to be gradually transmitted through resource allocation, operating activities, and accounting matters, and are ultimately reflected in financial indicators such as profitability, growth capability, and financial risk.

In combination with the preceding analysis, this study holds that BYD mainly affects corporate financial performance through the industrial-chain integration strategy, technological innovation strategy, and international expansion strategy. The specific transmission paths are as follows.

5.5.1 The Industrial-chain Integration Strategy Affects Profitability Through Cost Control

BYD has long adhered to the industrial-chain integration strategy. Through the layout of core links such as power batteries, automobile components, and complete-vehicle manufacturing, it improves the enterprise's control over supply-chain resources. Industrial-chain integration can reduce the enterprise's dependence on external suppliers, improve supply-chain stability, and reduce unit production costs through scaled production.

The transmission path is as follows: industrial-chain integration strategy → increased autonomy of core components → reduced production costs → formation of product cost advantages → improved profitability.

From the perspective of financial indicators, BYD's gross margin increased from 13.02% in 2021 to 19.44% in 2024, indicating that the industrial-chain synergy effect and scale advantages effectively improved the enterprise's profitability during the rapid development stage of NEVs. However, gross margin declined to 17.74% in 2025, showing that as competition in the NEV industry intensified, price competition weakened the effect of transforming cost advantages into profit advantages. Therefore, although the industrial-chain integration strategy constitutes an important competitive advantage for BYD, its promoting effect on profitability is still affected by changes in the industry competition environment.

5.5.2 The Technological Innovation Strategy Affects Corporate Growth Capability Through Achievement Transformation

The NEV industry has technology-intensive characteristics, and technological innovation capability is an important basis for an enterprise to maintain long-term competitive advantages. BYD has continued to increase R&D investment and has formed certain technological advantages in power batteries, hybrid technology, and core NEV technologies.

The transmission path is as follows: technological innovation strategy → increased R&D investment → formation of technological achievements → improvement of product competitiveness → expansion of market scale → enhancement of corporate growth capability.

R&D investment can improve an enterprise's long-term competitive capability, but because technological R&D has characteristics such as a long investment cycle and a complex achievement-transformation process, its economic value is not immediately reflected in financial indicators [14]. From the preceding analysis, BYD's R&D investment was about RMB 54.161 billion in 2024 and further increased to about RMB 63.441 billion in 2025, reflecting the enterprise's development strategy of continuously strengthening technological advantages. However, the growth rate of operating revenue declined during the same period, indicating that R&D investment needs to be further transformed through the commercialization of technological achievements and market competitive advantages before sustained financial returns can be formed. Therefore, when

© 2026 The Author(s). Authors retain copyright in their articles.

evaluating the effect of the technological innovation strategy, attention should not be paid only to the scale of R&D investment; the efficiency of R&D achievement transformation also needs attention [15].

5.5.3 The International Expansion Strategy Affects Financial Structure Through Capital Investment

As competition in the NEV industry gradually intensifies, BYD actively promotes overseas market layout and expands new growth space by increasing exports and building overseas production bases. International expansion can improve the enterprise's market coverage, but it also requires large amounts of capital for capacity construction, channel construction, and market development [16].

The transmission path is as follows: international expansion strategy → increased overseas investment → expansion of fixed assets and capital investment→ increased financing demand → influence on the enterprise's asset structure and solvency.

From the perspective of financial performance, BYD's asset scale has continued to expand in recent years, and fixed-asset investment has grown significantly, reflecting that the enterprise supports capacity expansion and global layout through capital investment. At the same time, the scale of liabilities has also increased with business expansion. Although the asset-liability ratio has declined in recent years, attention still needs to be paid to the efficiency of fund use and changes in capital structure during rapid expansion. Therefore, the internationalization strategy can not only provide new revenue growth space for the enterprise, but may also increase short-term financial pressure [17].

In summary, the process by which BYD's strategic factors affect financial performance can be summarized as follows: strategic choices affect the enterprise's resource allocation and operating activities, further act on accounting matters, and are ultimately reflected in changes in financial indicators. Among them, the industrial-chain integration strategy mainly affects profitability through the cost-control mechanism, the technological innovation strategy mainly affects growth capability through the achievement-transformation mechanism, and the international expansion strategy mainly affects financial risk through the capital-investment mechanism. It can be seen that after the NEV industry enters the stage of deepening competition from the stage of high-speed growth, enterprises need to pay attention not only to strategic advantages themselves, but also to the efficiency with which strategic advantages are transformed into financial results, so as to achieve long-term sustainable development.

6. Prospect Analysis

6.1 Development Prospects

The NEV market has broad development prospects. With the enhancement of environmental awareness, governments in various countries and regions have continuously introduced NEV subsidy policies to encourage NEV production. In addition, with the continuous development of science and technology, NEVs are also moving toward greater intelligence. In the long run, NEVs are expected to replace traditional fuel vehicles, and more intense market competition will follow [18]. As a leading enterprise in the NEV market, BYD's development prospects are worth expecting.

In terms of capacity and capital, because BYD has developed for a long time, it already has a complete industrial chain, complete industrial layout, and strong capacity. It has connected upstream and downstream suppliers, achieved self-sufficiency in some components, has abundant capital reserves, smooth capital-chain circulation, and relatively strong financing ability.

In terms of technological R&D, BYD has mature technological R&D capabilities, a large number of professional technical talents and large-scale scientific research teams, and at the same time holds many technological patents.

In terms of market expansion, BYD actively develops overseas markets, further expands global market share, continuously increases export volume, and has successfully covered 60 sovereign markets on six continents. In 2025, BYD's passenger-car exports reached 1.01 million vehicles, accounting for 55% of the increase in China's passenger-car exports, showing extremely strong growth momentum and global

© 2026 The Author(s). Authors retain copyright in their articles.

competitiveness. In addition, BYD has achieved significant results in market transformation. The launch of Denza and Yangwang marks BYD's breakthrough in the high-end market.

6.2 Development Suggestions

6.2.1 Improve the Efficiency of Industrial-chain Integration

The industrial-chain integration strategy is an important competitive advantage that distinguishes BYD from some NEV enterprises. It helps reduce the risk of external supply-chain fluctuations and improve the enterprise's production efficiency. However, as the NEV industry gradually enters the stage of deepening competition and market price competition continues to intensify, the cost advantage formed solely by relying on scale expansion is being challenged. From the perspective of financial indicators, BYD's gross margin and net profit margin both declined in 2025, indicating that the enterprise needs to further improve the utilization efficiency of industrial-chain resources and enhance the ability to transform strategic advantages into profitability. Therefore, in the future, BYD should shift from "expanding the scale of the industrial chain" to "improving the efficiency of the industrial chain." On the one hand, it should further optimize supply-chain management, improve the coordination efficiency of key component production and complete-vehicle manufacturing, and reduce production costs through digital management and lean production. On the other hand, it should pay attention to the utilization efficiency of new capacity, strengthen management of returns on fixed-asset investment, and avoid a decline in asset utilization efficiency caused by excessively rapid capacity expansion. At the same time, the enterprise also needs to improve product added value through technological upgrading and product-structure optimization, reducing the impact of price competition on profitability.

6.2.2 Improve the Transformation Efficiency of R&D Investment

The core of competition in the NEV industry is gradually shifting from scale competition to technology competition, and continuous R&D investment is an important basis for an enterprise to maintain long-term competitive advantages. In recent years, BYD has continuously increased R&D investment and has formed certain technological advantages in power batteries, hybrid technology, intelligent technology, and other fields, providing support for market expansion and product competitiveness improvement. However, R&D investment has characteristics such as a long investment cycle and a complex achievement-transformation process, and the expansion of R&D scale does not mean that the enterprise can immediately obtain financial returns. According to the preceding analysis, BYD's R&D investment continued to grow, but its operating revenue growth rate declined in 2025. This shows that the enterprise needs to further improve the allocation efficiency of R&D resources and transform technological advantages into operating results more effectively. Therefore, in the future, BYD should further optimize the structure of R&D investment and improve the commercialization efficiency of R&D achievements. On the one hand, it should strengthen continuous investment in core technology fields and maintain technological leading advantages in power batteries, intelligent driving, and other fields. On the other hand, it should establish a linkage mechanism between R&D investment and market demand, accelerate the transformation of technological achievements into product applications, and improve the contribution of R&D investment to revenue growth and profitability improvement. At the same time, in terms of the accounting treatment of R&D expenditure, continuous attention should also be paid to the impact of changes in the capitalization ratio on the comparability of financial information, so as to ensure that financial information can truly reflect the enterprise's innovation capability.

6.2.3 Optimize the Pace of International Expansion

As competition in the domestic NEV market intensifies, BYD actively promotes its internationalization strategy and expands new growth space by increasing overseas sales and building overseas production bases. Internationalized layout can help the enterprise break through the growth limitation of the domestic market and improve long-term growth capability, but overseas market development is also accompanied by relatively high demand for capital investment, including production-base construction, channel construction, brand promotion investment, and other inputs. From the preceding analysis, BYD's asset scale has continued to expand in recent years, fixed-asset investment has increased, and the scale of liabilities has also grown, indicating that international expansion requires a large amount of financial support. Therefore, in the process of promoting global development, the enterprise needs to balance growth targets and financial stability, avoiding pressure on cash flow and solvency caused by rapid expansion. In the future, BYD should adopt

© 2026 The Author(s). Authors retain copyright in their articles.

differentiated internationalization strategies according to the characteristics of different regional markets. In mature markets, it should pay more attention to brand building and product-structure optimization, and improve market recognition and profitability. In emerging markets, it should reasonably arrange production capacity in combination with local market demand and the industrial environment, so as to avoid overinvestment. At the same time, the enterprise should strengthen the evaluation mechanism for overseas investment projects and dynamically assess the effect of overseas expansion through indicators such as return on investment, cash flow, and asset utilization efficiency, so as to realize coordinated development between the internationalization strategy and financial-risk control.

7. Conclusion and Outlook

7.1 Conclusion

Based on the HAF and taking BYD as the research object, this study conducts research from four dimensions: strategic analysis, accounting analysis, financial analysis, and prospect analysis. On the basis of the HAF, this study further focuses on the process by which corporate strategic advantages are transformed into financial performance, and analyzes how strategic factors affect corporate financial performance through operating activities and accounting matters.

The study finds that the competitive advantages formed by BYD in recent years mainly come from strategic choices such as industrial-chain integration, technological innovation, and internationalized layout. Among them, the industrial-chain integration strategy strengthens the enterprise's control over power batteries, core components, and complete-vehicle manufacturing links, and improves profitability by reducing supply-chain costs and improving production efficiency. Financial data show that from 2021 to 2024, BYD's gross margin continued to increase, indicating that the industrial-chain synergy effect and scale advantages were effectively transformed into profit advantages during the high-speed development stage of the industry. However, as competition in the NEV market intensified, the enterprise's gross margin and net profit margin declined in 2025, indicating that the process of transforming strategic advantages into profit advantages was affected by industry price competition.

The technological innovation strategy is an important basis for BYD to maintain long-term competitive capability. The enterprise has continuously increased R&D investment and has formed technological advantages in power batteries, hybrid technology, and intelligent fields. However, R&D investment does not immediately transform into financial results, but needs to go through the process of technological achievement transformation and market application. Therefore, the enterprise needs to pay attention to the matching relationship between the scale of R&D investment and the efficiency of achievement transformation, so as to avoid evaluating strategic effects only through the growth of R&D investment.

The international expansion strategy provides BYD with new growth space, but it also brings capital-investment and financial-management pressure. Overseas market development, production-base construction, and channel investment all require large amounts of financial support and further affect the enterprise's asset structure and solvency. Therefore, in the process of promoting global layout, it is necessary to balance the speed of strategic expansion and the ability to control financial risks.

Overall, corporate strategic advantages do not directly form financial performance, but need to go through the transmission process of "strategic choice→ resource allocation → operating activities → accounting matters → financial indicators." By constructing the relationship between strategic factors and financial performance, this study further explains that the financial evaluation of NEV enterprises needs to focus not only on result indicators, but also to analyze the strategic logic behind the formation of financial results.

7.2 Outlook

Because this study mainly conducts analysis based on BYD's public financial data and case materials, there are still certain limitations. On the one hand, public data mainly reflect the enterprise's external operating results and cannot fully obtain micro-level information such as internal R&D management, supply-chain coordination efficiency, and overseas operating conditions. Therefore, the analysis of the specific action process of strategic factors still has certain limitations. On the other hand, this study takes a single enterprise

© 2026 The Author(s). Authors retain copyright in their articles.

case as the research object. Although it can deeply analyze the process by which BYD's strategic advantages are transformed into financial performance, the applicability of the research conclusions to different NEV enterprises still needs further verification. Future research can expand the sample range and explore the differences in the impact of different strategic models on financial performance through comparative analysis of enterprises in the same industry.

In addition, as the competitive environment of the NEV industry continues to change, the formation and transformation process of corporate strategic advantages also has dynamic characteristics. Future research can combine enterprise value-assessment models, financial forecasting models, and empirical analysis methods to further verify the mechanism through which strategic factors affect financial performance. At the same time, it can also combine factors such as digital transformation, the development of intelligent driving technology, and adjustment of global industrial chains to further expand the application scope of the HAF in NEV enterprise research. In addition, although the HAF has broken through the limitations of traditional financial analysis, there is still room for expansion, and it can be combined across disciplines to extend its application boundaries [19].

References

[1] Chen, A., Sha, Z., & Wen, S. (2023). Industry-chain support and the development of China's new energy vehicle industry. Fujian Tribune (Humanities and Social Sciences), (12), 54-72.

[2] Chen, H. (2019). Environmental analysis of new energy vehicles based on the PEST analysis method: Taking BYD as an example. Liaoning Economy, (5), 38-39.

[3] Zhang, Y. (2023). A comparative study of the international development of BYD and Tesla under the background of carbon neutrality (Master's thesis, Guizhou University of Finance and Economics).

[4] Jiang, L. (2016). Financial analysis and application of BYD based on the Harvard Analysis Framework (Master's thesis, Southwest Jiaotong University).

[5] Zhang, Z. (2025). Research on the competitive strategy of new energy vehicle enterprises: Taking BYD as an example. China Journal of Commerce, 34(5),156-159.

[6] Barney, J. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99-120. https://doi.org/10.1177/014920639101700108

[7] Qian, Y., Xu, G., Han, P., Sun, J., Wang, Y., & Shao, J. (2014). Vertical integration and internal capital allocation efficiency: A case study based on BYD. Accounting and Economics Research, 28(5), 41-54. https://doi.org/10.16314/j.cnki.31-2074/f.2014.05.001

[8] Wu, J., & Liang, L. (2023). Analysis of the financial status of new energy vehicles under the Harvard Analysis Framework: Taking BYD as an example. International Business Accounting, (14), 58-64.

[9] Yang, Z. (2013). Correlation research on the impact of R&D investment on enterprise value: An empirical test based on the implementation of new accounting standards by listed companies in China. Science and Technology Management Research, 33(10), 42-45.

[10] Ma, Z. (2020). Financial analysis of BYD Company based on the Harvard Analysis Framework (Master's thesis, Fujian Agriculture and Forestry University).

[11] Xia, L., Tao, Y., & Xin, Q. (2024). The impact of registration-system reform on accounting policy choices of IPO companies: Taking R&D expenditure capitalization as an example. Contemporary Accounting Review, 1(1), 1-19.

[12] Li, Y., & Hao, E. (2023). Financial analysis of BYD Company based on the Harvard Analysis Framework. Commercial Exhibition Economy, (4), 165-168.

[13] Lu, J., & Bian, P. (2025). Financial analysis of BYD Company. Co-Operative Economy & Science, (12), 141-145.

[14] Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), 128-152. https://doi.org/10.2307/2393553

© 2026 The Author(s). Authors retain copyright in their articles.

[15] Wu, X. (2025). Analysis of the impact of R&D investment on corporate financial performance: Taking BYD Company Limited as an example. Modern Industrial Economy and Informationization, 15(3), 227-228.

[16] Johanson, J., & Vahlne, J. E. (2017). The internationalization process of the firm—a model of knowledge development and increasing foreign market commitments. In International business (pp. 145-154). Routledge. https://doi.org/10.1057/palgrave.jibs.8490676

[17] Wang, C., & Wang, X. (2026). Analysis of the impact of the internationalization strategy of new energy vehicle enterprises on corporate financial performance from a dynamic perspective: Taking BYD as an example. Financial Management Research, (3), 52-64.

[18] Huang, Z., Hu, C., & Jin, J. (2026). Financial analysis of BYD Company based on the Harvard Analysis Framework. Business & Economy, (2), 161-164, 171.

[19] Xie, M. (2021). Financial analysis of BYD Company Limited based on the Harvard Analysis Framework (Master's thesis, Jiangsu University of Science and Technology).

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/).

© 2026 The Author(s). Authors retain copyright in their articles.