Asia Pacific Economic and Management Review

Editor-in-Chief:

Assoc. Prof. XuanDong Zhang

Yan'an University School of Economics and Management, China

Copyright © 2026. ASIA PACIFIC SCIENCE PUBLICATIONS COMPANY LIMITED. Complimentary Copy.

Asia Pacific Science Press

ISSN（Print）: 3005-9267
ISSN（Online）: 3005-9275

Asia Pacific Science Press

Asia Pacific Economic and Management Review

Asia Pacific Economic and Management Review (APEMR) is an international, peer-reviewed and open access journal which focuses on theoretical and applied studies of corporate and financial behavior. Aiming to promote the research in fields of business economics and management, it covers mainly but not limits to the following areas:

- Accounting and Financial Management
- Economics
- Political Economy
- Human Resource Management and Organizational Behavior
- Information Management
- International Business
- Strategy and Innovation
- Management Science and Operations Management
- Marketing and Retail
- Asset Evaluation
- Financial Theory and Practice
- Commercial Circulation
- International Trade
- E-commerce
- Logistics Management
- Business Administration

About Publisher:

Asia Pacific Science Press (APSP) is a swiftly expanding publisher of peer-reviewed and open-access journals, strategically located in Hong Kong. As a reliable and esteemed corporation, APSP is dedicated to promoting and serving a wide array of subject areas, ultimately contributing to the betterment of humanity. By disseminating knowledge to a global community of scholars, practitioners, researchers, and students, we strive to establish ourselves as the world's leading independent academic and professional publisher.

Submission instructions: You can submit your manuscript through the official website (www.apspublisher.com) or email (editor.apemr@apspublisher.com) , All manuscripts will go through a rapid peer review and production, making the process of publishing simpler and more efficient.

Publisher Headquarter

Room 03, 7th Floor, Block B, Tuen Mun Industrial Centre, 2 New Ping Street, Tuen Mun, Hong Kong, China
Website : www.apspublisher.com
Email : info@apspublisher.com

Fujian Province Office, China

603-1, 6th Floor, Building B20, Chengyi North Street, Software Park, Jimei District, Xiamen City, Fujian Province, China
Website : https://ojs.apspublisher.com/index.php/amit
Email : amit@apspublisher.com

Vol. 3 No. 3 (2026)

Asia Pacific Science Press

Table of Contents

1 **Research on Maintenance Cost Management of an Expressway Project**
Yanlong Zhang

16 **Analyst Attention, Institutional Investor Sentiment and IPO Pricing Efficiency**
Jinmei Zhao

27 **Research on the Strategic Conception and Realization Path of the Construction of ' Western Hydrogen Energy Financial Trading Center ' in Chengdu-Chongqing Economic Zone**
Pan Luo, Jinglong Kang

44 **The Impact of Green Finance on Regional Total Factor Productivity**
Qisong Wang, Shangjun Li, Yajie Han

62 **Research on the Impact of Enterprise Digitalization on the Governance of Multi-Partner R&D Alliances**
Jinlai Chen, Jie Liang, Shuhan Meng

71 **To Be or Not to Be? The Influence of Ambidextrous Leadership Behaviour on Ambidextrous Innovation: A Moderated Mediation Model**
Qian Gao

86 **Impact Analysis of the Financial Shared Services Model on Mengniu Dairy's Financial Data**
Hu Bo, Wang Ani

91 **Research on the Optimization of University Financial Management Models Empowered by Artificial Intelligence**
Fei Chen

97 **Research on the Enhancement of Technological Transformation Efficiency under New Quality Production Forces**
Wendi Cui

103 **Measurement and Evaluation of Regional Competitiveness of the Nine Provinces in the Yellow River Basin Based on Entropy Weight-TOPSIS Method**
Bingxin Wang

116 **ESG Integration and Long-Term Portfolio Performance: Evidence from International Markets**
Jiale Cai

121 **Job Crafting Under the Intervention of Large Language Models: Skill Premium, Task Evolution, and Professional Identity Conflicts**
Yijing Chen

126 **Construction and Application of Financial Early Warning Model for High-Tech Enterprises in the Era of Big Data**
Chuanjia Ma

132 **Analysis of the Influencing Factors of the Quality of Environmental Accounting Information Disclosure of Manufacturing Enterprises under the "Double Carbon" Goal**
Jiali Zhou

138 **Research on the Optimization of Professional Master Enrollment Promotion Model Based on Lasswell's "5W" Model**
Yang Wang, Pengzhi Kong

147 **The Digital Economy "Empowers" Carbon Productivity—Empirical Evidence from Chinese Cities**
Yu Zhang

158 **Research on the Inheritance Path of Intangible Cultural Heritage under the Strategy of Cultural Digitalization —Taking a Number of Intangible Cultural Heritage in Shandong Province as an Example**
Yaning Ren, Xiaoteng Huang

164 **Innovative Exploration of Cross-Campus Department Collaborative Governance System in Large Grade a Tertiary Hospitals from an Integrated Perspective**
Yiping Zhao

169 **Research on the Blended Learning Model for the Course of Bridge Engineering Construction Cost Planning and Management in the Context of Big Data——Using Plateau Complex Bridge Engineering as an Example**
Xiaoyan Liu, Yinheng Luo, Jinrong Yuan, Xiaolan Wang, Jiangjie Qin, Yao Li

174 **Research on Exchange Rate Risk Management in Corporate Foreign Currency Debt Financing: A Case Study of Evergrande Group**
Luyao Liang

180 **A Comparison of Global Aging Trends: Exploring the Path from "Healthy Aging" to "Active Aging" - Case Studies Based on Japan, Northern Europe and China**
Jiarui Hu, Mengyuan Fang, Jiahui Jiang, Haojun Huan, Zhengxiu Feng

184 **The Impact of Market-Incentivized Environmental Regulation on Green Innovation in Energy Enterprises**
Yi Ge

192 **Hybrid Selection and Adaptive Strategies with Clustering Strategies for Solving Multi-Objective Optimization Problems**
Manyuan Li, Meng Wen

205 **A Cross-Country Empirical Study on the Impact of Trade Openness on Economic Growth: A Comparative Analysis Based on Panel Data of 97 Countries**
Yuhan Yang, Shunxiang Bao

214 **CEO Compensation Contract Design under Moral Hazard and Adverse Selection**
Chenlu Fang

230 **A Study on the Spatio-Temporal Evolution and Determinants of Green Development Efficiency in the Cultural Industry of the Yangtze River Economic Belt**
Can Qin, Xuemei Zhang, Xiyue Yang, Rui Zhou, Yutong Cai

247 **Supply Chain Concentration and Corporate Default Risk: Evidence from China**
Renyuan Gu, Yangyue Huang, Qiying Li, Zihong Zhong, Qisong Wang

258 **Substation Indoor Environment Monitoring System based on STM32**
Lei Mei, Sixiao Zhang, Rui Mei, Yulong Shi, Ji Ma

267 **The Internationalisation Challenges of Chinese SMEs: Formation Mechanisms and Governance Pathways**
Hongyu Cao, Weixiang Gan

279 **The Impact of Women's Leadership on Firm Performance**
Shujun Chen

290 **A Critical Observation of SpaceX's Founding Phase**
Ziyao Yang

298 **Mapping Legitimacy Research in Business and Management: A CiteSpace-Based Bibliometric Review of Conceptual Diffusion and Thematic Evolution**
Zerong Huang, Doudou Zi, Tao Wang

314 **Research on the Multidimensional Network Spatial Structure and Correlation Effect of Urban Agglomerations : Analysis of China's Five Major Urban Agglomerations**
Tingying Peng, Xiaoyue Huang

330 **Research on Audit Agents Based on RPA Technology: A Case Study of Deloitte's Financial Robot**
Yanan Wei, Xuandong Zhang

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on Maintenance Cost Management of an Expressway Project

Yanlong Zhang*

Shandong Taidong Transportation Energy Construction Co., Ltd., Jinan, Shandong, 271100, China

Corresponding author: Yanlong Zhang, 594396043@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: With the vigorous advancement of expressway construction, China now boasts extensive highway coverage nationwide. However, most expressways have entered maintenance phases, making it imperative to balance high-quality transportation services with effective cost control in highway maintenance. This study focuses on the an Expressway Project, analyzing current maintenance cost control practices and identifying existing challenges. Through systematic evaluation of maintenance design, material procurement, construction implementation, and inventory management processes, we propose optimized cost control strategies. The research provides methodological guidance and theoretical support for cost management in A Expressway projects, enhancing their competitive edge and sustainable development. Additionally, the analytical framework developed for A Expressway maintenance cost control is expected to offer valuable insights for other expressway maintenance cost control models.

Keywords: Expressway Project; Maintenance Cost; Material Procurement; Construction Implementation

Published: Apr 28, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1321

1. Introduction

In recent years, with the rapid development of China's economy, the construction of expressway networks has made significant progress. By 2024, the total length of expressways in China had reached 183,600 kilometers, ranking first in the world. Expressways have significantly improved transportation capacity, optimized the road grade structure, and offer advantages such as speed, convenience, safety, and strong traffic flow capacity.

With the passage of time and increasing mileage, most expressways across China have entered maintenance phases. Notably, highways constructed during the 1980s and 1990s are gradually aging, prompting a shift from traditional minor repairs to comprehensive maintenance strategies. M. Hagood proposed that (2014)

Compared with preventive maintenance, rehabilitation, and major maintenance, routine maintenance is characterized by periodicity, repetitiveness, and timeliness (Li, 2021). Prolonged operational lifespans, natural erosion, human-caused accidents, and surging traffic volumes have exacerbated damage to highway infrastructure, significantly raising maintenance challenges. The entrenched "construction over maintenance" mindset has led to neglected repairs or inadequate maintenance practices, resulting in substantially higher long-term maintenance costs. Given that proper maintenance directly impacts highway longevity, travel service quality, and safety performance, maintenance efforts have become increasingly critical. With the continuous improvement of national road networks and growing vehicle populations, the demand for highway

maintenance services continues to rise.

To meet growing maintenance demands, the government has introduced multiple policies to guide the highway maintenance industry toward market-oriented, standardized, and intelligent development. Additionally, maintenance investment accounts for an increasing proportion of total highway investments, reflecting the market's urgent need for high-quality, in-depth maintenance services. MOTC (2006) conducted Taiwan has a total roadway lane length over 20,000 km, of which approximately 350 million square meters annually needs to be maintained and repaired for road surface and subgrade base. Kumar (2021) proposed that enterprises can effectively address pavement defect risks during maintenance operations through rational workforce allocation, efficient machinery deployment, and reliable material supply, thereby minimizing road asset losses and achieving optimal cost control. Scholar Bao (2021) conducted focused research on Pavement Condition Index (PCI) analysis. Liu (2022) conducted in-depth studies on optimizing maintenance budget utilization throughout highway operation cycles. Li Bin (2021) highlighted critical issues in highway project cost control such as lack of standardization, inaccurate cost estimation, incomplete supply chain management, unclear strategic positioning, insufficient cost analysis, and weak internal control mechanisms, with detailed elaboration on cost control standardization. Liu Fang (2021) found that most engineering bidding processes commonly suffer from issues such as inadequate management systems, particularly in non-public bidding categories. He proposed a comprehensive cost control accounting method for maintenance costs to achieve full-process cost management from planning, bidding, implementation, to maintenance, thereby realizing the goal of overall cost control. Scholar Yang Long (2021) combined the current operational realities of China's highways to conduct an in-depth analysis of the shortcomings in highway maintenance cost management and proposed several practical maintenance cost management methods. Lü Xinjian (2022), through case studies of highway maintenance across multiple provinces, argued that adopting the concept and methodology of full-process cost control and accounting for maintenance construction could effectively address issues such as inconsistencies between planned and actual progress, accelerated growth trends in maintenance costs, slow measurement and payment progress due to delayed acceptance, loopholes in safety production management, and high risks associated with maintenance funding.

Against this backdrop, this study takes the A Expressway project as a research case study. By analyzing maintenance cost data and management practices from 2019 to 2023, it identifies pain points and challenges in cost control across key stages. Targeted optimization strategies are developed for four critical areas: maintenance design, material procurement, construction implementation, and inventory management. The research aims to expand the scope and depth of cost control measures, transform traditional management models, maximize maintenance fund utilization efficiency, enhance the market competitiveness and operational sustainability of the A Expressway project, while providing theoretical insights and practical references for cost management in other domestic expressway projects.

2. Current Status of Maintenance Cost Control for Highway A Project

The A Expressway features a mainline spanning 200 kilometers and three connecting routes totaling 39.8 kilometers, with six lanes in each direction and a design speed of 120 km/h. As a key component of the Ministry of Transport's "Five Vertical and Seven Horizontal" national trunk road network, the A Expressway forms an interconnected transportation system that connects eastern and western regions while ensuring comprehensive connectivity. This study systematically analyzes maintenance cost control characteristics through five dimensions—overall costs, construction implementation costs, procurement costs, design costs, and period expenses—by examining maintenance cost data from 2019 to 2023.

2.1 Current Status of Maintenance Cost Control for Highway Project A

From 2019 to 2023, the total maintenance costs for the A Expressway Project increased from 909.9249 million yuan to 1,096.3992 million yuan, maintaining an annual growth trend. The primary maintenance costs rose from 415.5627 million yuan to 550.7213 million yuan, accounting for 50.23% of total costs compared to 45.67% previously, becoming the core cost component. Although total operating revenue grew from 1,554.1078 million yuan to 1,786.2889 million yuan, the significant increase in maintenance costs has exerted pressure on operational efficiency, highlighting the urgency of cost control. Table 1 presents the relevant maintenance cost data for the A Expressway Project during the 2019-2023 period.

Table 1. Maintenance-related costs for the A Expressway project Unit: 10,000 yuan

project	2019	2020	2021	2022	2023
Main maintenance costs	41,556.27	44,060.31	45,603.21	47,885.43	55,072.13
Other costs	1,237.72	1,270.35	1,259.66	1,345.72	1,296.83
total cost	90,992.49	95,409.94	105,270.57	105,544.29	109,639.92
Proportion of main maintenance costs to total costs	45.67%	46.18%	43.32%	45.37%	50.23%
Total operating revenue	155,410.78	161,381.91	168,457.11	173,763.51	178,628.89

As a transportation service industry, highway maintenance involves production costs that constitute a significant proportion of total maintenance expenses. These primarily include direct materials, labor costs, machinery expenses, and raw material procurement costs, as detailed below:

Construction process costs. The implementation of maintenance construction is a critical component of cost control for the A Expressway. The costs incurred during maintenance construction on the A Expressway are influenced by various factors, such as labor costs, material costs, and machinery expenses.

Table 2. Implementation Cost Table for Maintenance Construction of Highway Project A Unit: Ten Thousand Yuan

project	2019	2020	2021	2022	2023
direct labor	6,536.23	7,643.88	9,816.27	10,627.41	12,129.34
direct material	31,638.23	33,927.16	34,562.73	34,686.23	36,783.77
Direct machinery	2,473.16	2,538.96	9,837.16	1,782.52	1,227.38

As shown in the table above, labor costs for maintenance operations on Highway A exhibit a consistent annual upward trend. This is primarily attributed to the prolonged operation period of the highway, deteriorating road conditions, and inadequate preventive maintenance measures. Compared to previous years, the required workforce for repairing equivalent engineering volumes has increased significantly. Additionally, the highway's internal maintenance capacity remains insufficient, necessitating substantial external labor recruitment while wages continue to rise annually. Concurrently, regulatory oversight from highway police and road administration departments has become increasingly stringent. Construction or repair crews often receive immediate evacuation orders from traffic authorities or highway police before completing tasks due to heightened security requirements or heavy traffic volumes. This results in repeated repair operations requiring additional personnel. When construction is eventually permitted again, maintenance teams are forced to deploy large-scale manpower for emergency repairs. Although this approach fulfills dispatch requirements, it substantially increases labor costs.

Raw material procurement costs. The maintenance raw material procurement costs for the A Expressway project primarily include raw material procurement costs and employee compensation for the procurement department. The cost composition of the procurement phase for the A Expressway project maintenance from 2019 to 2023 is shown in Table 3.

Table 3. Composition of Maintenance Procurement Costs for Highway Project A Unit: 10,000 yuan

	2019	2020	2021	2022	2023
raw material procurement cost	31,234.73	32,089.12	33,983.25	34,267.91	36,326.79
employee compensation	139.21	143.66	145.08	150.17	152.63

The table above reveals that raw material procurement costs constituted the largest share of maintenance procurement expenses for the an Expressway Project between 2019 and 2023. The inherent nature of the transportation industry ensures sustained demand for raw materials throughout service operations. Procurement costs significantly influence maintenance expenditures. Key materials for A Expressway maintenance include routine maintenance supplies such as expansion joints, asphalt, barbed wire mesh, and guardrail panels; miscellaneous materials like pesticides for landscaping; bulk de-icing agents for snow removal and traffic flow maintenance; along with specialized maintenance materials used in various engineering

applications. Currently, all raw materials for A Expressway maintenance are sourced externally, making procurement costs highly susceptible to fluctuations in material prices.

Maintenance design serves as the foundation and prerequisite for the entire maintenance process, with all subsequent operations conducted based on this blueprint. The rationality of maintenance design directly determines the quality, efficiency, and effectiveness of post-maintenance tasks, thereby influencing overall cost control. Moreover, the quality of maintenance design significantly impacts the travel experience of drivers and passengers, which in turn exerts a decisive influence on the operational revenue of Highway A. Consequently, maintenance design constitutes a critical foundational task. The leadership of Highway A places high priority on maintenance design, allocating substantial resources for design expenditures as detailed in Table 4.

Table 4. Cost Composition of Maintenance Design Expenses for Highway Project A Unit: Ten Thousand Yuan

	2019	2020	2021	2022	2023
salary	3,737.92	3,805.81	5,437.13	5,506.68	5,614.56
travel expense	1,010.25	1,016.34	1,520.72	1,526.19	1,535.12
Outsourced design fee	5,291.80	5,339.81	8,415.98	8,586.19	8,701.56
other	386.26	453.19	661.98	633.1	653.04

The data in the table above reveals that wages and outsourced design fees accounted for the largest proportion of maintenance design costs in the A Expressway project from 2019 to 2023, collectively exceeding 80% of total design expenses. This proportion has shown a significant upward trend since 2021. The primary reason lies in the fact that after 2021, many outsourcing agencies shifted industries due to macroeconomic impacts, while the remaining few capitalized on the situation to continuously raise design fees. Additionally, the A Expressway project lacked effective internal exploration of maintenance design potential, coupled with insufficient employee awareness and capability in independent design, forcing reliance on external contractors. If this trend persists, maintenance design costs will become a substantial burden, while employee salaries continue to rise, adversely affecting overall cost control. This issue demands urgent attention.

Table 5. A: Cost Statistics During Maintenance Period of Expressway Projects Unit: Ten Thousand Yuan

project	2019	2020	2021	2022	2023
general expenses	61736.35	79324.81	81829.26	89826.11	98378.02
cost of financing	85205.11	82912.34	73178.29	65934.56	53126.72
inventory carrying cost	36723.72	38123.61	40731.85	43149.26	45932.71

The maintenance costs of Highway Project A primarily consist of administrative expenses, financial expenditures, and inventory management costs. As shown in Table 5, the administrative costs for highway maintenance during 2019-2023 exhibited a consistent annual increase, reflecting the Group Company's commitment to self-driven development strategies. Conversely, financial maintenance costs demonstrated a steady decline throughout this period, with a 37.65% reduction in 2023 compared to 2019. This improvement stems from the post-enterprise restructuring reforms, which enabled Highway Project A to adopt an independent operation model with self-accountability for profits and losses. Through systematic optimization of maintenance cost management systems, expenses such as service fees, late payment penalties, and interest charges have been progressively reduced. Meanwhile, inventory costs saw a gradual rise from 2019 to 2023, driven by rapid traffic volume growth that intensified maintenance demands and increased procurement of materials. Consequently, associated costs including inter-section transportation fees, storage charges, and warehouse rental expenses have also risen accordingly.

2.2 Maintenance Cost Control Measures for Highway Project A

Prior to initiating highway maintenance projects, conduct cost budgeting by gathering foundational data including road conditions, facility status, and traffic volume. Analyze maintenance requirements and cost drivers by integrating historical maintenance records with future planning strategies. Break down costs into categories such as routine maintenance,

engineering works, and equipment procurement to estimate labor, material, and equipment expenses. Compile the maintenance cost budget and allocate funds rationally, specifying cost targets for individual projects, materials, and labor costs. This ensures budget accuracy and operational feasibility of maintenance cost planning.

During maintenance operations, we strictly adhere to budgetary guidelines by conducting comprehensive cost accounting. All incurred expenses—including material costs, labor expenses, and equipment machinery costs—are accurately recorded and documented. We systematically collect and verify original vouchers, maintain accounting records, and regularly summarize and analyze maintenance cost ledger data to assess actual expenditure patterns while ensuring data authenticity and integrity. Additionally, we continuously monitor variance between actual costs and budget allocations, making timely adjustments to maintenance strategies to achieve effective cost control.

After maintenance operations are completed, conduct an in-depth analysis of the cost formation process and influencing factors to identify reasons for cost overruns or savings, thereby exploring cost reduction strategies. Simultaneously, evaluate cost control effectiveness to ensure accurate assessment of cost forecasting, decision-making, planning, control, and analysis processes, providing references for future maintenance cost management. Based on cost analysis and evaluation results, implement corresponding cost control measures such as optimizing maintenance workflows, reducing material consumption rates, and improving equipment utilization rates to minimize maintenance costs.

3. Issues in Maintenance Cost Control for the A Expressway Project

Through analyzing the current maintenance cost status and data of the an Expressway Project, combined with its actual operation and management processes, it was found that cost control issues primarily concentrate in four key areas: maintenance design, material procurement, construction implementation, and inventory management.

3.1 Unreasonable Maintenance Design

From conceptual design to final implementation, the entire process was independently managed by the Business Development Department, with minimal participation from related departments including Financial Management Department, Cost Accounting and Control Center, Maintenance Management Department, and Material & Equipment Management Center. When developing maintenance plans, the Business Development Department relied solely on historical experience and maintenance data analysis results. The final plans were formulated from the perspective of individual business units, focusing solely on analyzing maintenance requirements, priorities, and key areas while failing to comprehensively consider factors from other departments. This resulted in maintenance content, methods, timelines, and budgets that did not fully align with actual operational needs.

Furthermore, maintenance staff at the an Expressway Project were influenced by the “guaranteed income regardless of performance” mentality prevalent in public institutions before their corporate restructuring. This led to widespread low work motivation, with some employees adopting complacent attitudes and even exhibiting passive resignation behaviors. Despite consecutive annual salary increases, many workers perceived their current living standards as secure without additional effort, settling for status quo conditions and lacking ambition or competitive drive. Within this “guaranteed income” environment, employees showed insufficient attention to corporate maintenance and operational costs, subconsciously believing personal earnings were unrelated to cost control measures or business performance. The absence of cost-consciousness awareness among staff often resulted in severe resource wastage and reduced work efficiency.

3.2 Significant Procurement Risks for Materials

The procurement strategy for highway maintenance projects typically involves sporadic and decentralized purchases based on raw material requirements. This approach not only incurs time costs due to excessive procurement frequency but also risks mismatching supply quantities with actual demands. Lacking procurement autonomy, contractors often passively endure issues like suppliers’ sudden price hikes and delivery delays. Such disruptions can adversely affect construction schedules, exacerbate structural deterioration, compromise traffic safety, and escalate repair costs, ultimately leading to a sharp increase in overall expenses.

For maintenance operations of Highway A, materials with high daily demand and relatively low costs—including barbed wire mesh, isolation fences, guardrails, anti-glare nets, asphalt, cement, and sand/gravel—alongside technically demanding and

expensive components like expansion joints, directly impact overall maintenance cost control. Fragmented procurement practices, characterized by small order quantities and extended cycles, remain vulnerable to market price fluctuations, hindering cost management. Currently, Highway A primarily employs quotation and price comparison methods to control procurement costs, often selecting the lowest bid. However, purchasing departments frequently prioritize price over critical factors such as technical specifications, after-sales service, and contract fulfillment capabilities, resulting in costly overinvestment where premium materials are inadequately procured. Additionally, repetitive procurement approval processes consume excessive time, significantly compromising maintenance efficiency.

The maintenance department of Highway Project A has yet to establish strategic partnerships with suppliers in procurement. Currently, the procurement process primarily involves posting tender notices on the official website, allowing any supplier to submit bids. Coupled with the procurement department's reliance on methods such as price inquiries, this results in varying winning bidders each time, leading to discrepancies in the quality and specifications of supplied materials and equipment. Additionally, the lack of comprehensive process control often results in only post-event monitoring, creating numerous uncertainties throughout procurement and maintenance operations. Furthermore, frequent supplier changes and short collaboration cycles hinder effective communication between Highway Project A and suppliers, eroding trust foundations. Any temporary adjustments to procurement quantities or product categories require renegotiation and new contract signings, significantly extending procurement cycles and compromising cost control during subsequent maintenance phases.

3.3 Coarse Cost Control in Construction Implementation

During the construction of Highway A, the construction plan has proven insufficient in constraining cost control during implementation. The plan merely specifies requirements for machinery types and quantities, labor quotas, and maintenance cycles without addressing practical maintenance cost control needs through detailed operational plans for each construction phase. This has resulted in inadequate resource conservation awareness and low efficiency among frontline maintenance personnel. Maintenance sub-centers and teams mechanically follow work orders, as shown in Table 6 which outlines preventive maintenance data from recent years. Due to insufficient emphasis on preventive maintenance, minor defects often go untreated, escalating into major issues that require additional manpower, materials, and equipment for repairs. Consequently, cost control during actual maintenance operations remains challenging.

Table 6. A: Statistical Table of Preventive Maintenance Proportion for Highway Projects Unit: Number, Percentage

	Maintenance Task Order	Maintenance Task Order	Proportion of preventive maintenance
2019	23875	2141	8.97
2020	27967	2518	9.01
2021	33826	3689	10.91
2022	38641	3935	10.18
2023	40982	4683	11.43

The A Expressway exhibits deficiencies in safety production management during maintenance operations. From 2019 to 2023, substantial safety compensation expenditures were incurred, with detailed figures presented in Table 7. Consecutive workplace accidents in 2022 and 2023 not only resulted in disciplinary actions against responsible parties but also forced the expressway to bear significant post-incident remediation costs, with compensation amounts far exceeding market pricing benchmarks. The failure to utilize insurance reimbursement mechanisms exacerbated project cost overruns. In June 2022, improper placement of reflective cones in construction transition zones led to penalties from highway traffic police, causing seven-day work stoppages across the entire route, direct economic losses, and extended maintenance timelines. Additionally, maintenance crews—particularly subcontracted workers—demonstrated complacency and inadequate safety awareness. Daily operations frequently involved violations such as improper helmet usage, compromised safety protocols, failure to conduct mandatory three-minute pre-job briefings and safety training, and non-compliant operational procedures, resulting in workplace injuries, machinery damage, and other incidents. These systemic issues have created substantial challenges for cost

control in A Expressway construction projects.

Table 7. A: Statistical Table of Maintenance Safety Compensation Fees for Expressway Projects Unit: Ten Thousand Yuan

project	2019	2020	2021	2022	2023
security indemnity	372.93	581.13	943.88	1,393.71	1,811.85

3.4 Inadequate Inventory Management Measures

Currently, warehouses at grassroots units such as the Maintenance Sub-center and maintenance teams of Highway A project remain inadequately stocked, resulting in daily maintenance materials not being stored according to regulations. Raw materials required for maintenance operations are stored independently by grassroots units. Analysis from Table 8 shows that raw material inventories at Highway A increased annually from 2021 to 2023 with significant growth rates. The average raw material expenditure reached 111.4482 million yuan, while inventory write-offs also showed an upward trend. On one hand, inventory management for Highway A's maintenance operations primarily relies on grassroots units, where warehouse managers handle stock entry and exit procedures. However, challenges including numerous grassroots units, aging warehouse managers, and low educational backgrounds have slowed the implementation of standardized inventory management systems. Insufficient warehouse facilities at most grassroots units have led to suboptimal maintenance environments, disorganized material storage, and even instances of raw materials being mixed with scrap materials. On the other hand, chaotic inventory management practices—including excessive stockpiles without proper sorting and scattered storage locations—have not only increased company inventory verification costs but also caused time wastage, disrupted construction schedules, and indirectly elevated maintenance expenses.

Table 8. Statistical Table of Maintenance Inventory for Highway Project A Unit: Ten Thousand Yuan

	2019	2020	2021	2022	2023
Raw material inventory	9,630.19	9,912.76	10,616.29	10,983.27	11,834.91
inventory balance	19,837.45	20,351.16	22,853.82	23,926.17	24,947.15
amortization amount	212.13	267.14	327.86	361.93	398.52

4. Optimization Scheme for Highway Cost Control

To address cost control challenges in A Expressway maintenance projects, this study proposes targeted optimization strategies covering four critical phases: maintenance design, material procurement, construction execution, and inventory management. These solutions aim to drive comprehensive management upgrades across all operational stages.

4.1 Optimization Scheme for Maintenance Design Cost Control

An expressway should break down overarching objectives into specific tasks and targets for each department, ensuring clear understanding of their roles and contributions to corporate goals. For instance, when designing maintenance plans, departments such as Financial Management, Maintenance Management, Cost Accounting & Control Center, and Materials & Equipment Management Center should collaborate to fulfill their respective responsibilities, avoiding the Development Department taking on sole responsibility. Maintenance Management Departments, Sub-centers, and Work Teams must actively provide raw data to support subsequent maintenance design. Simultaneously, an information communication mechanism should be established through regular or ad-hoc multi-department meetings to facilitate joint problem-solving, information sharing, and resource integration. Leveraging internal networks and platforms like DingTalk management systems, we should develop efficient information-sharing platforms to ensure seamless communication channels. Critical progress updates and dynamic data should be readily accessible, enabling timely aggregation of key information for decision-makers and management to make informed decisions. Additionally, regular workflow optimization should be implemented to ensure smooth inter-departmental coordination. Clear collaboration guidelines and systems must be established to provide operational direction. Through team-building activities and training programs, employees' teamwork spirit should be

strengthened, helping staff truly understand that unity and cooperation are essential for achieving mutual growth between individuals and the organization.

Enterprises should regularly organize cost awareness education programs, utilizing training sessions and seminars to educate employees on the importance and fundamental methods of cost control, thereby enhancing staff cost consciousness. Concurrently, establish a performance evaluation mechanism linked to cost performance, integrating cost control metrics into employee assessment systems and correlating compensation, promotions, and other incentives with cost control outcomes to motivate active participation. Implement comprehensive budget management by incorporating maintenance expenditures for Highway A into budgetary frameworks, strictly controlling waste during execution to ensure effective cost containment. Develop a cost control accountability system that leverages internal staff potential, boosts motivation, strengthens cost awareness, and fulfills assigned tasks while reducing outsourcing expenses. Clearly define responsibilities and obligations across departments and positions to foster a culture of collective cost control participation. Establish robust cost control protocols, rewarding departments demonstrating exceptional performance while holding underperforming units accountable. This incentive-driven approach enhances employee engagement and ensures effective implementation of cost control measures. Cultivate an organizational culture emphasizing cost management practices, enabling all staff to naturally adopt cost-conscious operational philosophies in daily operations.

4.2 Optimization Plan for Material Procurement Cost Control

Highway authorities should proactively implement cross-departmental coordination mechanisms to enhance procurement process flexibility. By conducting in-depth analysis of operational conditions across units and forecasting future procurement demands, they can utilize data analytics tools such as historical purchase records and market trends to improve prediction accuracy. Establishing a supplier evaluation and selection system will ensure the recruitment of high-quality, reliable partners. Implementing quality control procedures guarantees that procured materials meet established standards. Real-time procurement progress monitoring provides critical data for subsequent maintenance planning. Regular training programs should be organized to equip procurement staff with up-to-date knowledge, skills, and regulatory compliance training. Inviting industry experts to share practical insights will continuously elevate professionals' expertise. Encouraging open communication channels between procurement teams and other departments enables timely understanding of grassroots-level procurement needs. By systematically analyzing procurement patterns at the operational level, procurement personnel can strengthen collaboration capabilities while gaining comprehensive insights into authentic requirements across departments and frontline units.

The Maintenance Management Department of Highway Project A should comprehensively integrate procurement requirements for labor, materials, and equipment through centralized bidding processes to establish bulk purchase orders. This approach significantly enhances bargaining power with suppliers while reducing fragmented procurement approvals, cutting processing time, and markedly improving procurement efficiency. Departments may organize joint meetings for strategic planning, conduct quantity verification against annual maintenance construction blueprints, and accurately calculate procurement demands for labor, materials, machinery, and equipment. This enables targeted division of construction sections. By referencing Highway A's historical operational data, annual bidding plans should be formulated with proactive procurement strategies. Consolidating similar procurement items and expanding centralized procurement scale will strengthen scientific rigor, timeliness, foresight, and accuracy in procurement operations, ultimately achieving comprehensive efficiency improvements.

The Maintenance Management Department of Highway Project A should maintain close communication with suppliers through regular meetings, phone calls, and emails. This ensures timely sharing of critical factors affecting maintenance efficiency, including material/equipment requirements, quality standards, and delivery locations, thereby guaranteeing seamless information flow. Relevant highway departments must conduct periodic comprehensive evaluations of suppliers' performance across quality metrics, delivery compliance, cost-effectiveness, and service support capabilities. When maintenance budgets permit, proactive measures should be taken to provide essential training and technical assistance to enhance suppliers' management expertise and technical proficiency. Both parties should jointly develop contingency plans to facilitate prompt

problem-solving through coordinated communication, minimizing time losses caused by interdepartmental coordination delays.

4.3 Optimization Plan for Construction Implementation Cost Control

Based on differences in service scope and content, operation centers are categorized into three types: material operation centers, mechanical equipment operation centers, and labor personnel operation centers. For detailed information, refer to Table 9 below.

Table 9. Analysis of Operation Centers and Driving Factors for Maintenance Installation Panel Projects in Highway A Expressway Project

Work Center	school assignment	Activity driver
Material Operations Center	Guardrail panel	Guardrail panel tonnage
	lean concrete	Concrete volume
	tiling	Number of square bricks per ton
	upright column	Number of column roots
Mechanical Leasing Operations Center	Crane Rental	Number of cranes
	Excavator rental	Excavator count
	Pile Driver Rental	Number of pile drivers
	Early Warning Vehicle Rental	Number of warning vehicles
	sporadic machinery leasing	Number of scattered machinery
Labor Service Center	Guardrail panel team operation	Guardrail panel installation
	Bricklayer team operation	Concrete and Masonry Construction
	Closing team operations	Work area cleaning

Through in-depth analysis and meticulous consideration of resource drivers at each activity center, the types and quantities of resources utilized by each activity center are clarified. The specific analysis results of resource drivers for each activity center are presented in Table 10.

Table 10. A Analysis of Operation Centers and Resource Drivers for Guardrail Installation Maintenance Project at the A Highway Project

Work Center	school assignment	Activity driver
Material Operations Center	Guardrail panel	Labor Day
	lean concrete	Labor Day
	tiling	Labor Day
	upright column	Labor Day
Mechanical Leasing Operations Center	Crane Rental	Labor Day
	Excavator rental	Labor Day
	Pile Driver Rental	Labor Day
	Early Warning Vehicle Rental	Labor Day
	sporadic machinery leasing	Labor Day
Labor Service Center	Guardrail panel team operation	Labor Day
	Bricklayer team operation	Labor Day
	Closing team operations	Labor Day

Through in-depth analysis of operational activities and related resource drivers, we consolidated resource costs generated across all activity centers to establish the project activity library cost structure, while systematically summarizing cost profiles

for each center. Material costs, machinery equipment rental expenses, and labor costs associated with this project are detailed in Table 11, Table 12, and Table 13 respectively.

Table 11. Material Costs for Guardrail Installation Maintenance Project of Highway A

construction stage	material	unit	quantities	unit-price	Total price (RMB)
Main construction	Guardrail panel	t	1035	1578	1633230
	lean concrete	m^3	560	380	212800
	tiling	m^3	165	95	15675
	cast-in-place shoulder stone	m^3	81	105	8505
	upright column	root	2070	367	759690
Installation and Construction	Anti-glare facilities	Xiang	1	158432	158432
	Contour material	m^2	500	249	124535
	Traffic sign material	m^2	368	162	58960
	Other materials	Xiang	1	8950	8950
amount to					2980777

Table 12. A: Machinery and Equipment Rental Costs for Guardrail Installation Maintenance Project of Highway Expressway

construction machinery	unit	quantity	Lease unit price (RMB/month)	Lease term (months)	Lease cost (RMB)
backhoe	tower	6	8000	2	96000
crane	tower	3	5500	2	33000
monkey engine	tower	6	5000	2	60000
Early warning vehicle	moon	6	3000	2	36000
Other machinery	Xiang	1	10000	2	20000
amount to					245000

Table 13. Labor Costs for Guardrail Installation Maintenance Project of Highway A

Assignment Project	primary coverage	Total price (RMB)
traditional cure	Excavation and backfilling	180000
Guardrail panel	Transportation, installation, and calibration	550000
concrete bed	Cast-in-place, curing, installation	98300
Signboard production	install	86000
pouring	Transportation, pouring, and curing	60760
road	Construction of Road shoulders and curbs	30000
Other assignments	Other construction	36000
amount to		1041060

Table 14. Labor Costs for Guardrail Installation Maintenance Project of Highway A

Fee Name	the cost of the project	Cost (RMB)
Salary for team leaders	Salary + Performance Bonus	83200
car fare	Travel expenses + meal allowance	13000
Other transaction costs	Office expenses + Labor insurance premium	3670
amount to		99870

The maintenance and installation of guardrail panels for the A Highway project involves numerous costs during the transaction process, including project information fees and business management fees, among others. These expenses are challenging to control and estimate. However, by employing statistical analysis methods to examine these transaction costs, stakeholders can achieve more precise cost management and control. As demonstrated in Table 15.

Table 15. Transaction Costs for Highway Maintenance Project: Guardrail Panel Installation

Project Information Fee	Business management fee	amount to
8000	90000	98000

Regarding the maintenance and installation of guardrail panels for the A Expressway project, Table 16 details other direct costs, Table 17 outlines indirect costs, and Table 18 shows the breakdown of period expenses for the enterprise.

Table 16. Other Direct Costs for Highway Maintenance and Guardrail Installation Project on Highway A

Cost item	the cost of the project	charge against revenue
Garbage disposal fee	$650m^2$ × 34 yuan/sqm	22100
Security Measures Fee	2% of the contract amount	300000
Construction water and electricity charges	Water and electricity charges	8000
Large machinery entry and exit fees	Entry and Exit Fees for Large Machinery Equipment	62000
amount to		392100

Table 17. Indirect Costs of Guardrail Installation Maintenance Project for Highway A

Cost item	the cost of the project	charge against revenue
Project staff salaries	Project staff salaries	831000
travel expense	Transportation costs + subsidies	22170
charges	Pollution discharge fee + social security fee, etc.	16820
Other indirect costs	Business entertainment expenses, etc.	54000
amount to		923990

Table 18. Periodic Cost of Enterprise for Highway Project Maintenance Installation of Guardrail Panels on Highway A

Cost item	the cost of the project	charge against revenue
performance bond processing fee	Amount of performance bond × 0.125%	18750
insurance	All Risks Insurance, Safety Liability Insurance, Work Injury Insurance	45000
general and administrative expenses	Corporate management fee payable	113220
Safety Production Expense	Safety design, facility layout	84830
expenses of taxation	engineering tax	1050000
amount to		1311800

According to the cost statistics of the A Expressway Project maintenance and installation of guardrail panels, the transaction costs incurred during the transaction period, including project personnel salaries and transportation expenses, amounted to 99,870 yuan.

Labor Service Center Cost Driver Allocation Rate:

1041060/（1041060+245000+2980777）=24.40%

Material Activity Center cost driver allocation rate:

2980777/（1041060+245000+2980777）=69.86%

Cost driver allocation rate for equipment leasing operations center:

245000/（1041060+245000+2980777）=5.74%

The transaction cost of the labor supply operation center is:

99870 × 24.40% = 24,368.28 yuan

The transaction cost of the material processing center is:

99870 × 69.86% = 69,769.18 yuan

The transaction cost of the equipment leasing operation center is:

The total cost for the guardrail installation maintenance project of Highway A (99870 × 5.74% = 5732.54 yuan) is detailed in Table 19.

Table 19. Total Cost of Guardrail Installation Maintenance Project for Highway A

Cost item	the cost of the project	charge against revenue
material cost	Civil construction, installation, and auxiliary materials	2980777
Equipment leasing cost	Mechanical equipment and small tools	245000
cost of service rendered	labor employment	1041060
Other direct costs		392100
project indirect cost		923990
interperiod cost		1311800
customer transaction cost		99870
amount to		6994597

Using guardrail panel installation as a case study, this paper demonstrates how Activity-Based Costing (ABC) is applied in corporate cost management accounting. Other operations can reference this analytical approach to obtain precise cost estimates for various activities.

(1) The calculation steps for guardrail panel labor costs are as follows:

In the guardrail installation project for Highway A maintenance, the labor contract signed with the main structural construction team explicitly stipulated centralized procurement for guardrail panels. Accordingly, the labor team's responsibilities were limited to panel cutting and installation. After comprehensive evaluation of all relevant factors, the transaction cost between the project and the labor supplier was determined to be 24,368.28 yuan.

The operational transaction cost for guardrail panel engineering is:

550,000 ÷ 104,1060 × 24,368.28 = 12,873.95 yuan

The total actual installation cost for guardrail panels is:

The total cost of 550,000 yuan plus 12,873.95 yuan equals 562,873.95 yuan. As shown in Table 11 (Material Cost Table), the total weight of installed guardrail panels is 1,035 tons, indicating the labor cost for installation as follows:

562,873.95 yuan per 1,035 tons = 543.84 yuan per ton

(2) The calculation steps for mechanical costs of installing guardrail panels are as follows:

In the installation project of guardrail panels, the mechanical equipment used is not directly related to the specifications and materials of the guardrail panels. Based on this, when calculating machinery costs, installation expenses can be determined according to different models of guardrail panels.

Table 20. A: Machinery Rental Costs for Highway Maintenance and Guardrail Installation Project

Cost item	unit	quantity	Usage time (months)	Rent (RMB)	Total cost (RMB)
cutting machine	tower	6	2	2000	24000
arc welder	tower	6	2	2000	24000
Other miscellaneous machinery	Xiang	1	2	1200	2400
amount to					50400

The calculation above shows that the transaction cost between the project and the machinery leasing supplier amounts to 5,732.54 yuan.

The transaction cost for installing guardrail panels and operating machinery rental is:

50400/245000 × 5732.54 = 1179.26 yuan

The actual cost for installing guardrail panels through mechanical leasing is:

50,400 + 1,179.26 = 51,579.26 yuan

As shown in Table 11 Material Cost Table, the total weight of installed guardrail panels is 1,035 tons, resulting in a per-ton installation machinery cost of:

51579.26 ÷ 1035 = 49.83 yuan per ton

(3) The material cost calculation steps for centralized procurement of guardrail panels are as follows:

By integrating project and material supplier transaction cost allocation mechanisms, the cost driver allocation rate for guardrail panel installation projects at material operation centers is as follows:

1633230/2980777=54.79%

The transaction cost for guardrail panel material procurement is:

69,769.18 × 54.79% = 38,226.53 yuan

The actual procurement cost of guardrail panels during installation is as follows:

1633230 + 38226.53 = 1,671,456.53 yuan

As shown in Table 11 Material Cost Table, the total installation volume of guardrail panels is 1,035 tons, resulting in a centralized procurement cost of [amount] per ton.

1671456.53 ÷ 1035 = 1614.93 yuan per ton

(4) The calculation steps for other direct costs associated with guardrail panel installation are as follows:

Other direct costs associated with guardrail panel installation include inspection and testing fees, primarily assessing structural integrity, surface quality, dimensions, raw materials, anti-corrosion coatings, and toughness. The standard fee is 80 yuan per set. Based on the actual on-site testing costs (including sample submission fees) of 8,850 yuan, the actual additional direct costs incurred per ton of guardrail panels are calculated as follows:

8850/1035 = 8.6 yuan per ton

(5) The cost accounting steps for guardrail panel installation are as follows:

The above calculations indicate that the direct cost for the guardrail installation maintenance project of Highway A is as follows:

550,000 + 50,400 + 1,633,230 + 885 = 2,242,480 yuan

The total direct costs incurred during construction are as follows:

1633230 + 245000 + 2980777 + 392100 = 5251107 yuan

Therefore, the operational motivation rate for the guardrail installation maintenance work under the A Expressway Project is as follows:

2242480/5251107=42.70%

The total of project indirect costs, administrative overheads, and transaction costs between the project and clients amounts to:

923990 + 1311800 + 98000 = 2,333,790 yuan

Therefore, the indirect costs incurred in the A Expressway project's guardrail installation maintenance work are as follows:

2333790 × 42.70% = 996528.33 yuan

The indirect cost per ton of guardrail panel installation is calculated as: 996,528.33 ÷ 1,035 = 962.83 yuan per ton.

In summary, the actual cost per ton for guardrail panel installation works is:

543.84 + 49.83 + 1614.93 + 8.6 + 962.83 = 3180.03 yuan

4.4 Optimization Plan for Inventory Management Cost Control

The construction of facilities such as warehouses not only reduces material and equipment loss rates but also enhances maintenance construction efficiency. Conduct comprehensive inspections of existing warehouses and implement customized

upgrade strategies for each facility to ensure rational and orderly material storage. Firstly, warehouse construction enables materials to be preserved according to maintenance management requirements, mitigating adverse effects caused by natural factors. Secondly, it facilitates inventory management through regular stocktaking and clearance processes, allowing timely identification and resolution of inventory management issues, reducing stockpiling, and improving inventory turnover rates. Thirdly, categorizing materials by type, properties, and quantity enables quick location retrieval during daily maintenance operations, saving time while ensuring accuracy in material classification and quantities to support construction activities. Fourthly, coordinated equipment scheduling revitalizes idle assets, expands equipment leasing services, and creates new profit growth opportunities.

Based on the locations of maintenance sub-centers and teams, strategically establish regional emergency material central warehouses. Collaborate with functional departments including Maintenance Management Department, Business Development Department, and Cost Accounting & Control Center to conduct rational procurement of materials and equipment according to historical demand patterns and current maintenance plans. These materials are then allocated to designated emergency material warehouses, enabling centralized procurement, tiered management, and unified distribution to effectively reduce material costs. When maintenance sub-centers or teams require replenishment or face stock shortages, resources can be promptly allocated from nearby emergency material warehouses, significantly shortening delivery times and lowering inventory management costs. Concurrently, enhance warehouse administrator training programs to improve professional skills and ethical standards. Select responsible and meticulous staff members for warehouse management positions, while implementing robust incentive mechanisms to boost employee motivation.

5. Conclusion

This study conducts an analysis and research on maintenance cost control for the A Expressway project, expanding the scope of cost management to achieve greater comprehensiveness. The objective is to provide innovative approaches for maintenance cost control in A Expressway projects. The research yields the following conclusions:

The A Expressway project should promptly transition from its current traditional maintenance cost control approach to expand both the scope and depth of cost management strategies. Traditional cost control methods primarily focus on internal factors during highway maintenance operations, overlooking external influences such as supplier relationships and client interactions. While these approaches may yield noticeable short-term results, they often fail to meet long-term cost control objectives. Therefore, adopting modern cost management models becomes imperative. Based on the current maintenance status of the A Expressway project, we have identified key challenges through research and implementation adjustments. By addressing critical aspects including maintenance design, material procurement, construction execution, and inventory management, we have systematically developed optimized cost control solutions.

Against the backdrop of continuously growing highway maintenance demands and escalating cost pressures nationwide, highway operators must abandon the outdated approach of prioritizing construction over maintenance and single-point control. By identifying cost management pain points based on operational realities, they should establish comprehensive, precision-driven maintenance cost control systems to drive industry transformation toward marketization, standardization, intelligentization, and cost efficiency. Future strategies should integrate smart technologies—including big data analytics, IoT solutions, and AI systems—across all maintenance processes. This integration will enable accurate demand forecasting, automated material procurement, real-time construction monitoring, and dynamic inventory management, ultimately achieving intelligent cost optimization and enhanced operational efficiency in highway maintenance.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Li, B. (2021). Cost control measures for expressway projects. Green Environmental Protection Building Materials, (6), 78–79.

[2] Liu, F. (2021). Cost control analysis of highway maintenance. Heilongjiang Transportation Science and Technology, 44(10), 205.

[3] Yang, L. (2021). Analysis and application of cost management control in expressway projects. Transport Manager World, (34), 61–63.

[4] Dong, X. M. (2021). Highway engineering project management and project cost control. Heilongjiang Transportation Science and Technology, 44(9), 233–234.

[5] Wei, J. L., Chen, H. L., & Lin, L. (2022). Construction of a smart management platform for highway maintenance materials. China Transportation Informatization, (S1), 137–139.

[6] Lu, X. J., Sun, C., & Dou, W. B. (2022). Exploration of the application of artificial intelligence in highway maintenance services. China Transportation Informatization, (S1), 452–453.

[7] Wang, X. P. (2022). Characteristics and issues analysis of expressway maintenance management. Transportation World, (21), 182–184.

[8] Liu, Y. (2022). Exploration on enterprise cost control and management based on value chain. China Small and Medium Enterprises, (5), 83–84.

[9] Zhang, G. (2023). How construction enterprises use value chain theory to control costs. China Business World, (9), 122–123.

[10] Zhu, X. P. (2022). Research on construction cost control of A expressway project (Master's thesis). Guizhou University.

[11] Shi, Y. Y. (2022). Cost management and control measures during the operation period of expressways. Yunnan Hydropower, 38(3), 168–170.

[12] Zhu, Y. (2024). Application of management accounting in corporate financial management. Cooperative Economy and Technology, (18), 133–135.

[13] Wang, B. W. (2022). Research on enterprise cost control and management based on value chain. Finance & Economics World, (27), 27–29.

[14] Zhu, Y. Y. (2024). Discussion on tax risk cost control strategies in corporate financial management. Accounting Learning, (5), 143–145.

[15] Yi, Y. L. (2023). Cost control research of Haier smart home based on value chain. Industrial Innovation Research, (6), 168–170.

[16] Ministry of Transportation and Communications. (2006). Ministry of Transportation and Communications, Taiwan. Retrieved November 2006, from http://www.thb.gov.tw/

[17] Hagood, M. (2014). Highway routine maintenance cost estimation for Nevada. University of Nevada, Las Vegas.

[18] Kumar, R., & Suman, S. K. (2021). International journal of systems assurance engineering and management pub. Behavioral Sciences, 39, 9–13.

[19] Bao, S. T., Han, K. Y., Zhang, L., Luo, X. D., & Chen, S. Q. (2021). Applied sciences pub. Highway Safety, 25, 10–18.

[20] Liu, S. S., Arifin, M. F. A., & Chen, W. T. (2022). IEEE access pub. Structural Cost Management Accounting Horizons, 33, 1–27.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Analyst Attention, Institutional Investor Sentiment and IPO Pricing Efficiency

Jinmei Zhao*

School of Management, University of Shanghai for Science and Technology, Shanghai, 200093, China

Corresponding author: Jinmei Zhao, 18757293895@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Against the backdrop of the continuous deepening of the registration-based reform and the emphasis on self-reliance and self-strengthening in science and technology in the 15th Five-Year Plan, the IPO pricing efficiency of the STAR Market, as a core board serving hard-tech enterprises, has become a key issue of concernin the capital market. This paper takes the listed companies on the STAR Market from 2019 to 2024 as the research sample and systematically examines the impact of analysts' concerns on the IPO pricing efficiency and its mechanism of action. Research has found that analysts' attention has reduced the efficiency of IPO pricing. Further research has revealed that analysts' attention, by enhancing the sentiment of institutional investors, has exacerbated the degree of IPO underpricing, verifying that analysts have a tendency to cater to institutional investors and amplify irrational market demands during the information transmission process. The research conclusions of this paper have certain theoretical significance and practical implications for regulating the professional conduct of analysts, improving the information disclosure system, and enhancing the IPO pricing efficiency under the registration system.

Keywords: Analyst Focus IPO Pricing Efficiency Institutional Investor Sentiment; Registration-Based Reform

Published: Apr 29, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1320

1. Introduction

The year 2026 marks the beginning of the 15th Five-Year Plan, and self-reliance and self-strengthening in science and technology have been established as a core national strategy. As a "test field" for serving scientific and technological innovation, the STAR Market has seen over 600 listed companies since its opening in 2019, becoming a gathering place for hard-tech enterprises. From the perspective of the market operation cycle, the current market is in the middle of the third stage of a complete upward cycle. Valuations are being reassessed, market divergence is increasing, and industry differentiation is significant. Sectors with high growth in performance will achieve excess returns. The technology sector has a significant performance advantage and may become a hot investment option. Meanwhile, the registration-based reform has been continuously deepened, the overall plan for the "1+6" reform of the STAR Market has been implemented, and the marketization degree of the IPO pricing mechanism has been constantly enhanced. Regulatory authorities have made it clear that institutional investors such as social security funds can participate in strategic placement, including institutional investors in the scope of strategic investors and strengthening their pricing power in new share issuance. Against this backdrop, how to enhance the pricing efficiency of IPOs on the STAR Market is an important issue that urgently needs to be addressed as

China's capital market gradually matures.

At present, the market pricing efficiency is mainly measured by the rate of return on issuance and listing (Rock, 1986), and the existing literature on pricing efficiency roughly starts from the following three aspects: The first aspect is how to solve the problem of "unclear value". Among them, analysts are important information intermediaries in the capital market and have a significant impact on market pricing. Analysts can significantly reduce the information friction between issuers and investors by unearth and "translate" the true value of a company through actions such as releasing research reports and conducting on-site investigations. Foreign scholars have pointed out that the coverage of analysts can enhance the reflection of IPO prices on the fundamentals of companies, thereby reducing the level of underpricing (Givoly and Lakonishok, 1979; Lys and Sohn, 1990). However, some studies have found that analysts colluded with institutional investors, leading to pricing deviations (Xu Nianxing et al., 2012; Shao Xinjian et al., 2018). Guan Feng et al. (2024) further discovered that the correlation between analysts and institutional investors is a key factor influencing the direction of their actions. The second aspect is to make "pricing more reasonable". Generally, from the institutional perspective, Guo Xuan et al. (2025), based on the pilot program of the registration-based reform of the Growth Enterprise Market, used the difference-in-differences method to study and found that the registration-based reform significantly improved the IPO pricing efficiency by enhancing the quality of information disclosure, verifying that high-quality information disclosure is a key factor in improving market pricing efficiency. The research by Lai Li et al. (2022) indicates that market-oriented reforms can enhance the pricing efficiency of the Chinese market, and the registration-based IPO reform has positive institutional effects. Wu Xihao and Zhang Chi (2022) found from the perspective of IPO underpricing rate that the registration-based reform has a significant impact on the pricing efficiency of the capital market. Xue Shuang and Wang Yu (2022) found through text analysis of responses to inquiry letters on the STAR Market that responses to inquiry letters can enhance the efficiency of IPO pricing, and their mechanism of action is related to the divergence in quotations among institutional investors. The third aspect is "how to influence pricing". Underwriter behavior, institutional and retail investor sentiment, corporate governance, and investor speculation all belong to important factors influencing pricing in the market. Among them, regarding the behavior of institutional investors, some scholars believe that institutions use their inquiry rights to make the issue price closer to the intrinsic value, and they also believe that an increase in the shareholding ratio of institutional investors can effectively play a supervisory and governance role and improve pricing efficiency (Gao et al. (2020). However, when the shareholding ratio is too high and there is a long-term business relationship with the underwriter, it is easy to have a supportive quote, believing that it will push up the issue price and increase the risk of long-term price breaking (Ljungqvist, A., Nanda, V., & Singh, R., 2006). The higher the proportion of institutional shareholding, the more significant the herd effect, which is prone to trigger buying high and selling low, leading to short-term overreaction in IPO pricing and reducing pricing efficiency (Gompers, P. A., & Metrick, A., 2001).

This paper mainly focuses on the first and third directions. Due to the differences in knowledge and information acquisition between investors and issuers, investors often find it difficult to fully and accurately assess the true value of IPO companies, resulting in pricing deviations. Under the traditional approval system, the issue price is often artificially depressed to attract investors to participate, resulting in a frequent sharp increase in the price of new stocks after their listing. Although the registration system has enhanced market transparency and analysts' attention has to some extent disclosed corporate information, information asymmetry and irrational market behavior still pose challenges to IPO pricing. Moreover, the impact of IPO underpricing by institutional investors in the primary market and individual investors in the secondary market is not the same. Therefore, research analysts' attention to the impact of institutional investors in the primary market on the efficiency of IPO pricing and its mechanism of action can not only provide valuable policy references for improving the quality of information disclosure in securities issuance, but also help reveal the deep mechanism of IPO pricing under the registration system, further enriching academic research on IPOs in China's stock market.

2. Theoretical Analysis and Research Hypotheses

2.1 Analyst Focus and IPO Pricing Efficiency

Under the registration system, the degree of information disclosure has gradually increased. Based on the value discovery hypothesis, analysts, through their professional information collection capabilities, can provide the market with incremental

information related to company value and reduce the degree of information asymmetry between investors and IPO companies (Givoly & Lakonishok, 1979; Lys & Sohn, 1990). In the context of IPOs where information asymmetry is particularly severe, companies listed on the STAR Market have characteristics such as short establishment time, novel business models, and a lack of comparable companies, making valuation and pricing more difficult (Guan Feng et al., 2024). At this point, the participation of more analysts means a greater supply of information and cross-validation, which helps to form more accurate price discoveries. Meanwhile, competition among analysts can also curb the extreme predictions of individual analysts, making market pricing more reasonable (Hong et al., 2000). Furthermore, analysts' attention can attract investors' attention, prompt more market participants to focus on IPO companies, and thereby enhance the information efficiency of market pricing (Bajo & Raimondo, 2017). However, based on the conflict of interest hypothesis, analysts are not completely independent information intermediaries, and their behaviors are influenced by multiple interest relationships (Mehran & Stulz, 2007; Ramnath et al., 2008). During the IPO issuance process, underwriters and issuers have a strong motivation to raise the issue price, while institutional investors hope to lower the issue price to obtain higher returns. The main service targets of analysts are institutional investors, and their income is directly linked to the transaction commission allocation of institutional investors and the New Fortune selection (Firth et al., 2013; Gu et al., 2013), thus there is a tendency to cater to the interests of institutional investors. Especially when analysts have alumni, same-location or commission relationships with institutional investors, they are more likely to issue biased forecast prices and collude with institutional investors to jointly depress the IPO issue price (Guan Feng et al., 2024). Shao Xinjian et al. (2018) found that analysts colluded to overestimate in IPO pricing. More analysts' attention means more possibilities of collusion, which may intensify the degree of IPO underpricing and undermine pricing efficiency.

H1: Under the condition that other factors remain unchanged, analyst attention is significantly positively correlated with the IPO underpricing rate, that is, the increase in analyst attention reduces the IPO pricing efficiency.

2.2 Analyst Focus, Institutional Investor Sentiment and IPO Pricing Efficiency

Under the institutional background of the registration-based reform in China's capital market, securities analysts, as key information intermediaries, are playing an increasingly important role in the IPO pricing process. Meanwhile, prospect theory holds that investors are more likely to prefer investing in lottery stocks with positive skewness in returns, just like the IPO stocks under the registration system implemented on the STAR Market: On the one hand, compared with listed stocks, a portion of the value of IPO stocks exists as call options, and their returns have a positive skewness, which is more attractive to investors (Green and Hwang, 2012; Chen and Zheng, 2021); On the other hand, compared with IPO stocks that have price fluctuation limits under the approval system, IPO stocks under the registration system have no price fluctuation limits in the first five days of listing, and their returns are more positively skewed, with a higher probability of achieving greater investment returns, which is more attractive to investors. In reality, the prevalence of IPO underpricing has led market investors to generally form a fixed perception that "subscribing to new shares guarantees profits", and analysts' attention has further magnified this market expectation, continuously stimulating investors to participate in new share subscriptions and secondary market transactions. First, as the core body of the offline allocation of the STAR Market, institutional investors, under the guidance of the optimistic value signals and high growth expectations continuously released by analysts, will not only significantly increase the offline subscription intensity in the face of the potential new share subscription returns of the STAR Market, but also make early arrangements for secondary market transactions to seize excess returns, thereby pushing up their own allocation ratio and optimistic sentiment. Meanwhile, most of the enterprises listed on the STAR Market are technology innovation companies with novel business models and high valuation difficulties. Institutional investors highly rely on professional research reports from analysts to judge their value. The higher the analysts' attention, the more concentrated the value information institutions can obtain, and their optimism also rises accordingly. Secondly, subsequent investors who enter the market will blindly follow the previous market trading heat and the trading behaviors of institutional and individual investors when making decisions on new share trading. This further promotes the continuous fermentation and rise of optimistic sentiment, ultimately driving the closing price of new shares on the first day of listing on the STAR Market to rise significantly, widening the gap between the issue price and the closing price, and intensifying the degree of IPO

underpricing. Based on the above analysis, this paper proposes the following assumptions:

H2: Under the condition that other factors remain unchanged, analysts focus on enhancing the sentiment of institutional investors, intensifying IPO underpricing, and reducing pricing efficiency.

3. Research Design

3.1 Sample Selection and Data Sources

This paper takes the listed companies on China's STAR Market from 2019 to 2024 as the initial research samples and conducts the following processing: (1) Exclude enterprises with abnormal trading such as ST and PT; (2) Exclude enterprises in the financial and real estate sectors; (3) Eliminate samples with missing values; (4) To control the influence of outliers, on the basis of eliminating missing values, the continuous variables were truncated by 1% above and below, and finally 567 pieces of observed data were obtained. The data is from the CSMAR database and Wind database.

3.2 Variable Selection and Model Design

3.2.1 Variable Selection

(1) IPO pricing efficiency (UP). When measuring the pricing efficiency of an IPO, in order to directly reflect investors' initial value perception bias of the enterprise and avoid long-term market noise interfering with the assessment of pricing efficiency, this paper refers to (Lai Li et al., 2022) and selects the IPO underpricing rate constructed based on the first-day window. The calculation formula is

$$\frac{P_1-P_0}{P_0}$$

and take the logarithm of it for processing, and P_0 is issue price, P_1 is closing price on the first day of listing.The larger the UP value is, the lower the issue price is compared to the market price on the first day, and the higher the degree of IPO underpricing.

(2) Analyst Focus Breadth (Analyst_Breadth). This article takes the IPO issuance date as the benchmark and counts the number of analysts who have released research reports within three months prior to the issuance date. The data is sourced from the Wind database.

(3) Institutional investor sentiment. This article refers to (Li Ke et al., 2023) and uses the shareholding ratio of institutional investors to measure the sentiment of institutional investors. The specific numerical calculation method is the ratio of the number of new shares allocated offline to the number of valid offline subscriptions.

(4) Control variables. The specific control variables are shown in Table 1.

Table 1. Variable Definition Table

Project	Variable name	Variable symbol	Variable definition
The explained variable	IPO pricing efficiency	UP	Ln (Closing price on the first day of listing/ Issue price-1)
Mediating variable	Institutional investor sentiment	Investors	The number of shares allocated through the offline placement of the IPO/The number of valid offline subscription shares×100%
Explanatory variable	The breadth of analysts' attention	Analyst_Breadth	The number of analysts paying attention to in the three months before an IPO
Control variable	Company scale	In_Size	The natural logarithm of total assets at the end of the year prior to listing (Unit: Ten thousand yuan)
	Retrun on net assets	ROE	Return on Net Assets for the Year Prior to Market Launch
	Whether it has been audited by the Big Four	Big4	Dummy variable: 1 if the auditor is one of the Big Four accounting firms; otherwise, 0

Project	Variable name	Variable symbol	Variable definition
	Asset-liability ratio	Leverage	The total liabilities at the end of the year prior to listing divided by the total assets at the end of the year
	Issuance pricing level	asinh_Pricing	Perform an inverse hyperbolic sine transformation on the diluted initial public offering price-earnings ratio
	Turnover rate on the first day of listing	Turnover	Trading volume on the first day of listing/Circulating share capital×100%
	Research and development investment	RD	Research and development expenses/Operating income
	The winning rate of online issuance	HitRate	The winning ratio for online IPO subscription
Virtual variable	Industry dummy variable	INDUSTRY	Set dummy variables for listed companies
	Year dummy variable	Year	Set dummy variables by year

3.2.2 Model Design

To test Hypothesis 1 of this paper, a regression model (1) is set up. Among them, the core explanatory variable is the number of analysts (Analyst_Breadth) who paid attention to the enterprises listed on the STAR Market in the three months prior to their IPOs.$CONTROL_M$ represent a collection of control variables，At the same time, the fixed effects of Year and INDUSTRY are added to the model. ε represents the random error term, and the standard error is adjusted by clustering at the enterprise level.

$$UP=\beta_0+\beta_1 Analyst_Breadth+\sum \beta_m CONTROL_M+\sum \beta_i INDUSTRY_i+\sum \beta_j Year_j+\varepsilon \quad (1)$$

Analysts are concerned that the mechanism of enhancing institutional investors' sentiment can lead to an increase in IPO underpricing rates and a decline in pricing efficiency. Referring to the research of Li Ke et al. (2023), this paper constructs model (2) for mechanism testing.

$$Investors_hold=\beta_0+\beta_1 Analyst_Breadth+\sum \beta_m CONTROL_M+\sum \beta_i INDUSTRY_i+\sum \beta_j Year_j+\varepsilon \quad (2)$$

3.3 Descriptive Statistics

Table 2 reports the descriptive statistical results of the main variables. The results show that during the sample period, the average value of analyst attention was 4.490, the median was 2.000, and the standard deviation was 5.860, indicating that on average, approximately 4.49 analysts followed the sample companies during the IPO period, but there were significant differences among different companies. This distribution characteristic is basically consistent with existing studies and also provides a basis for us to test the heterogeneous impact that analysts are concerned about. The average value of the IPO underpricing rate was 1.235, the median was 0.874, the minimum was -0.298, and the maximum was 6.598. This indicates that during the sample period, there was a common phenomenon of high underpricing on the first day of listing of new stocks. However, there was a significant differentiation among different companies, and some companies even experienced price drops below the issue price. This feature provides sufficient variation for studying the influencing factors of pricing efficiency. The average value of institutional investor sentiment is 16.963, the median is 11.369, the standard deviation is 15.976, and the maximum value is 73.902. This indicates that the average shareholding ratio of institutional investors after an IPO is approximately 17%, but there are significant differences among different companies. Some companies have attracted a large number of institutional investors, while others have received almost no interest from institutions. This distribution feature provides a good data basis for us to examine the mechanism effect of institutional investor sentiment between analyst attention and pricing efficiency.

Table 2. Descriptive Statistics of Main Variables

	N	Sum	Mean	SD	Min	Max	Median	p75
Analyst Breadth	569	2555.000	4.490	5.860	0.000	27.000	2.000	6.000
UP	569	702.675	1.235	1.321	-0.298	6.598	0.874	1.762
Investors hold	569	9651.890	16.963	15.976	0.013	73.902	11.369	26.118
HitRate	569	22.570	0.040	0.016	0.024	0.150	0.037	0.045
RD	569	26552.284	46.665	93.140	0.000	421.780	9.330	18.340
In Size	569	11835.553	20.801	0.985	19.183	24.381	20.586	21.302
ROE	569	535.628	0.941	9.810	-52.890	22.411	2.355	5.949
Big4	569	61.000	0.107	0.310	0.000	1.000	0.000	0.000
Leverage	569	180.666	0.318	0.197	0.042	0.800	0.280	0.462
Asinh Pricing	569	2380.991	4.185	1.523	0.000	6.841	4.453	4.888
Turnover	569	12.577	0.022	0.019	0.002	0.094	0.017	0.027

4. Empirical Analysis

4.1 Benchmark Regression Results

Based on Model (1), this paper conducts regression calculations, and the results are shown in Table 3. The results show that the coefficient of analyst attention is 0.023 in column (3), which is significant at the 5% level. This indicates that after controlling for other factors and unobstructed heterogeneity, the higher the analyst attention, the higher the IPO underpricing rate, that is, the lower the IPO pricing efficiency. This result supports research hypothesis H1.

Table 3. Analyst Concerns and IPO Underpricing

	(1)	(2)	(3)
	UP	UP	UP
Analyst_Breadth	0.018**	0.015	0.023**
	(0.008)	(0.011)	(0.009)
In_Size		-0.259***	-0.200***
		(0.075)	(0.067)
ROE		0.014***	0.009**
		(0.005)	(0.005)
Turnover		-6.828***	-1.873
		(2.374)	(2.199)
Big4		0.379**	0.298**
		(0.186)	(0.150)
Leverage		1.657***	1.061***
		(0.313)	(0.264)
asinh_Pricing		-0.100**	-0.052
		(0.043)	(0.037)
HitRate		-8.272	-0.442
		(5.074)	(4.338)
RD		-0.001**	-0.001*
		(0.001)	(0.001)
Constants	-0.183***	5.557***	3.872***
	(0.055)	(1.494)	(1.353)
Industry Fixed	No	No	Yes
Year Fixed	No	No	Yes
Observed value	567	567	567
Adjusted R^2	0.280	0.107	0.306

Note: ***, ** and * represent significance levels of 1%, 5% and 10% respectively, with T values in parentheses. All coefficients have been corrected for robust standard errors.

4.2 Mediating Effect

This paper refers to the mediating effect analysis method proposed by Jiang Ting (2022) to examine the specific mechanism of the impact of enterprise digital transformation on stock price synchronicity. The core idea of this mediating effect test is to first conduct a regression test to examine the causal relationship between the explanatory variable and the mediating variable, and then explain the influence of the mediating variable on the explained variable through existing theoretical and literature research conclusions, thereby determining whether a mediating effect exists. The results show that after controlling for industry and year fixed effects, the coefficient of analyst attention in column (3) is 1.024, which is significantly positive at the 1% level, indicating that the higher the analyst attention, the higher the sentiment of institutional investors.

Regarding the impact of institutional investor sentiment on the IPO underpricing rate, existing research has provided solid theoretical support. The "Ownership dispersion Hypothesis" holds that IPO underpricing increases the number of shareholders and ownership dispersion by attracting more investors to subscribe and leading to proportional allocation during the placement process. That is, the underpricing rate is positively correlated with the breadth of ownership (Bouzouita et al., 2015). From the perspective of corporate governance, it has been proved that the issuer intentionally attracts institutional investors to hold shares by underpricing in exchange for its supervisory function after listing. Underpricing is a rational choice to attract institutional holdings (Stoughton and Zechner, 1998). In domestic research, for companies with a higher proportion of institutional investors' holdings, the market pricing efficiency is more significantly affected by the quality of interaction, indirectly confirming the important role of institutional holdings in the pricing process (Bian Shibo et al., 2022). The above theories jointly indicate that there is an inherent positive correlation between the shareholding ratio of investors and the IPO underpricing rate. Therefore, it is theoretically valid for analysts to focus on the transmission mechanism that boosts the sentiment of institutional investors, thereby increasing the underpricing rate of IPOs and reducing pricing efficiency. Based on the empirical results in Table 3 and Table 4, the core hypothesis H2 is thus supported.

Table 4. Intermediary Role of Institutional Investors' Shareholding Ratio

	(1)	(2)	(3)
	Investors_hold	**Investors_hold**	**Investors_hold**
Analyst_Breadth	0.949***	0.954***	1.024***
	(0.124)	(0.136)	(0.122)
In_Size		0.386	-1.530**
		(0.771)	(0.774)
ROE		0.029	0.083
		(0.066)	(0.062)
Turnover		48.935	12.692
		(39.765)	(38.983)
Big4		-3.797*	-2.036
		(2.212)	(2.014)
Leverage		-3.199	3.021
		(3.669)	(3.901)
asinh_Pricing		0.938*	0.955**
		(0.493)	(0.463)
HitRate		-114.253***	-49.305
		(41.067)	(37.601)
RD		0.016**	0.005
		(0.007)	(0.007)
Constants	12.724***	4.817	40.867***
	(0.745)	(15.133)	(15.113)

	(1)	(2)	(3)
	Investors_hold	**Investors_hold**	**Investors_hold**
Industry Fixed	No	No	Yes
Year Fixed	No	No	Yes
Observed value	567	567	567
Adjusted R^2	0.190	0.130	0.212

Note: ***, ** and * represent significance levels of 1%, 5% and 10% respectively, with T values in parentheses. All coefficients have been corrected for robust standard errors.

4.3 Robustness Test

(1) Other methods for measuring IPO underpricing. The measurement method of IPO underpricing has a significant impact on the analysis and final conclusion of this paper. To ensure the robustness of the results, this paper redefines IPO underpricing as: (the closing price of the 30th trading day of the new share listing - the issue price of the new share)/the issue price of the new share. The regression results show that the regression coefficient in Table5 (1) is significantly positive, indicating that the higher the analyst's attention, the more significantly the underpricing of new ipos on the STAR Market increases. After changing the measurement method of IPO underpricing, the empirical conclusion still remains robust and reliable.

(2) Replace the explanatory variable. The number of people analysts focus on is usually right-skewed. To avoid affecting the analysis results of this paper, the logarithmic processing method will be adopted. The results are shown in Table 5 (2), and the regression coefficient is significantly positive. At the same time, a dummy variable Analyst_Dummy is constructed. If an analyst releases a research report or conducts a survey, the value of Analyst_Dummy is 1, indicating that there is an analyst's concern. Otherwise, the value is 0, indicating that no analysts are paying attention. This variable only focuses on "whether there is" analyst tracking to test the impact of the analyst's attention to this behavior itself on the IPO pricing efficiency. The results are shown in Table 5 (3), and the regression coefficients are significantly positive. The empirical conclusion is consistent with the above benchmark regression.

Table 5. Robustness test results

	(1)	(2)	(3)
	UP_30	**UP**	**UP**
Analyst_Breadth	0.027***		
	(0.003)		
Analyst_Ln		0.164***	
		(0.021)	
Analyst_Dummy			0.207***
			(0.045)
Constants	2.333***	2.178***	1.727***
	(0.526)	(0.530)	(0.536)
Industry Fixed	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes
Observed value	567	567	567
Adjusted R^2	0.407	0.410	0.363

(3) Endogeneity test. Regression analysis shows that analysts' attention can significantly increase the IPO underpricing rate, and the increase in the IPO underpricing rate may in turn affect analysts' tracking decisions by attracting more investors' attention and stimulating market enthusiasm. It can be seen that there may be an endogeneity problem of mutual causality between analysts' attention and IPO pricing efficiency. Therefore, to address the endogenous bias caused by bidirectional causality, it is necessary to first control the interactive fixed effects between industries and years. Data statistics show that there are significant differences among IPO companies in different industries and years in China in terms of market

environment, regulatory policies, etc. Therefore, a fixed interaction effect between industries and years is set and controlled, and then a benchmark regression is conducted. The results are shown in column (3) of Table 3. The estimated coefficient is positive and significant at the 5% level, suggesting that even after excluding the interference of industry and annual macro factors, analysts' attention still has a positive effect on the IPO underpricing rate. This result remains robust.

Meanwhile, to further alleviate the possible endogeneity problem, this paper adopts the two-stage least squares method for regression and selects the average analyst attention of other companies in the same industry - quarter (IV_Analyst) as the instrumental variable. The construction logic of this instrumental variable lies in the fact that companies in the same industry and quarter face similar market environments and industry prosperity, and their analysts' attention is highly correlated, meeting the correlation requirements of the instrumental variable. Meanwhile, the attention paid by analysts from other companies will not directly affect the IPO pricing efficiency of this company, meeting the exogenous requirement. The regression results show that the F-statistic of the first stage of regression is 17.22, which is higher than the critical value of 16.38. Therefore, there is no need to consider the problem of weak instrumental variables. The estimation coefficient of the two-stage least squares method is positive and significant at the 1% level, indicating that on the basis of considering the endogeneity issue, the increase in analysts' attention does indeed lead to an increase in the IPO underpricing rate, that is, to reduce the IPO pricing efficiency. In conclusion, after controlling for the endogeneity issue, the core conclusion of this paper still holds true.

Table 6. Endogeneity Test

	(1)	(2)
	Analyst_Breadth	**UP**
IV_Analyst	0.452*** (0.109)	
Analyst_Breadth		0.161*** (0.044)
Controlled variable	Yes	Yes
Industry Fixed	Yes	Yes
Year Fixed	Yes	Yes
Kleibergen-Paap rk LM		22.8
Kleibergen-Paap rk Wald F		17.22
R^2	0.324	0.082

5. Conclusions and Suggestions

This paper takes the listed companies on the STAR Market of China's A-share market from 2019 to 2024 as the research object and systematically examines the impact of analysts' attention on the efficiency of IPO pricing and its mechanism of action. Research findings indicate that analyst attention is significantly positively correlated with the IPO underpricing rate, meaning that an increase in analyst attention significantly reduces the IPO pricing efficiency. Mechanism tests show that analyst attention boosts the IPO underpricing rate by increasing the shareholding ratio of investors. The above findings support that analysts' attention can influence IPO pricing by attracting investors' attention and stimulating irrational demand, rather than merely playing the role of information transmission. Therefore, based on the above test results, this paper puts forward the following policy suggestions for better leveraging the information intermediary function of analysts and enhancing the efficiency of IPO pricing under the background of the registration system:

(1) Strengthen the performance norms of analysts and other intermediary institutions, and solidify their market responsibilities. This paper's research indicates that analysts' attention has not enhanced pricing efficiency as expected by traditional theories; instead, it may intensify market speculation by attracting attention. Regulatory authorities should further clarify the boundaries of analysts' responsibilities during the research and release process and enhance their sense of diligence

and responsibility. At the same time, through appropriate institutional arrangements, analysts should be urged to better perform their functions of information mining and value discovery, rather than releasing inflammatory predictions to cater to market sentiment.

(2) Enhance the authenticity, relevance and readability of information disclosure during the process of new share issuance, and guide the market to return to value-driven. The information disclosure system should be investor demand-oriented, encouraging listed companies to proactively disclose relevant information that affects their decisions from the perspective of investors, and achieving a shift in the focus of information disclosure from “supervisor-oriented” to “investor-oriented”. In addition, information disclosure should highlight differentiation, with stratified disclosure based on the characteristics of different industries and enterprises of different scales, to enhance the readability and decision-making usefulness of information, and effectively reduce the behavioral tendency of investors to overly rely on the attention of analysts due to a lack of information.

(3) Strengthen the education and protection of investors and gradually optimize the structure of market investors. Under the premise of maintaining market stability, we will continue to expand and optimize the institutional investor team, give full play to its function as a market stabilizer, guide long-term funds to enter the market through institutional design, gradually optimize the structure of market participants, and enable professional institutions to play a greater role in price discovery. In addition, customized investor education programs can be implemented to enhance investment knowledge education for individual investors. Regular publicity activities on rational investment and risk prevention should be carried out to guide them to establish a long-term investment concept and reduce blind pursuit of short-term hotspots.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Guo, X., Jiang, D. F., & Tao, L. (2025). Registration system reform, information disclosure quality and IPO pricing efficiency. Business Review, 11(20), 100–107, 115.

[2] Guan, F., Xu, Y. F., & He, K. G. (2024). The role of analysts in IPO market-based pricing: Evidence from STAR market companies. Accounting and Economic Studies, 38(2), 47–70. https://doi.org/10.16314/j.cnki.31-2074/f.20240703.001

[3] Li, K., Lin, Y. J., & Qi, B. (2023). Shares issued registration system reform, investor sentiment and IPO underpricing. Journal of Contemporary Finance and Economics, (10), 58–71. https://doi.org/10.13676/j.cnki.cn36-1030/f.20230915.004

[4] Wu, X. H., & Zhang, C. (2024). Research on the impact of the registration-based reform on the pricing efficiency of the capital market: From the perspective of IPO underpricing rate. Nankai Business Review, 27(2), 246–258.

[5] Jiang, T. (2022). Mediation effect and adjustment effect in the study of causal inference experience. China Industrial Economy, (5), 100–120. https://doi.org/10.19581/j.cnki.ciejournal.2022.05.005

[6] Bian, S. B., Chen, Y., & Wang, X. X. (2022). Can high-quality interaction enhance the pricing efficiency of the stock market? Based on the research of "the Shanghai e interaction". Journal of Economics (Quarterly), 22(3), 749–772. https://doi.org/10.13821/j.cnki.ceq.2022.03.02

[7] Lai, L., Lan, C. D., & Qin, M. C. (2022). Has market-oriented reform improved pricing efficiency? Evidence from registration system. Management World, 38(4), 172–184, 185–190, 199. https://doi.org/10.19744/j.cnki.11-1235/f.2022.0059

[8] Xue, S., & Wang, Y. (2022). Kechuang board IPO audit information reply back to the starting underpricing. Management World, 38(4), 185–203. https://doi.org/10.19744/j.cnki.11-1235/f.2022.0055

[9] Shao, X. J., Hong, J. J., & Liao, J. C. (2018). Analyst collusion overvaluation in China's new share issuance and its welfare impact. Economic Research Journal, 53(6), 82–96.

[10] Xu, N. X., Jiang, X. Y., Yi, Z. H., et al. (2012). Analyst conflict of interest, optimism bias and stock price crash risk.

Economic Research Journal, 47(7), 127–140.

[11] Rock, K. (1986). Why new issues are underpriced. Journal of Financial Economics, 15(1–2), 187–212.

[12] Weber, J., Willenborg, M., Wu, B., et al. (2023). IPO price formation and analyst coverage. Review of Accounting Studies, 1–50.

[13] Ljungqvist, A., Nanda, V., et al. (2006). Hot markets, investor sentiment, and IPO pricing. The Journal of Business, 79(4), 1667–1702.

[14] Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. The Quarterly Journal of Economics, 116(1), 229–259.

[15] Givoly, D., & Lakonishok, J. (1979). The information content of financial analysts' forecasts of earnings: Some evidence on semi-strong inefficiency. Journal of Accounting and Economics, 1(3), 165–185.

[16] Gao, S. H., Brockman, P., et al. (2020). Differences of opinion, institutional bids, and IPO underpricing. Journal of Corporate Finance, 60, 101540.

[17] Lys, T., & Sohn, S. (1990). The association between revisions of financial analysts' earnings forecasts and security-price changes. Journal of Accounting and Economics, 13(4), 341–363.

[18] Bajo, E., & Raimondo, C. (2017). Media sentiment and IPO underpricing. Journal of Corporate Finance, 46, 139–153.

[19] Hong, H., Lim, T., et al. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies.The Journal of Finance, 55(1), 265–295.

[20] Yang, N. J., Lv, Y. H., & Yang, H. L. (2023). Does media sentiment shape the underpricing of IPO in China? EAI.

[21] Mehran, H., & Stulz, R. M. (2006). The economics of conflicts of interest in financial institutions. Journal of Financial Economics, 85(2), 267–296.

[22] Ramnath, S., Rock, S., et al. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. International Journal of Forecasting, 24(1), 34–75.

[23] Firth, M., Lin, C., & Liu, P. (2013). The client is king: Do mutual fund relationships bias analyst recommendations? Journal of Accounting Research, 51(1), 165–200.

[24] Gu, Z. Y., Li, Z. Q., & Yang, Y. G. (2013). Monitors or predators: The influence of institutional investors on sell-side analysts. The Accounting Review, 88(1), 137–169.

[25] Chen, H. H., & Zheng, M. R. (2021). IPO underperformance and the idiosyncratic risk puzzle. Journal of Banking and Finance, 131.

[26] Green, T. C., & Hwang, B. H. (2012). Initial public offerings as lotteries: Skewness preference and first-day returns. Management Science, 58(2), 432–444.

[27] Bouzouita, N., & Gajewski, J. F. (2015). Liquidity benefits from IPO underpricing: Ownership dispersion or information effect. Financial Management, 44(4), 785–810.

[28] Stoughton, N. M., & Zechner, J. (1998). IPO-mechanisms, monitoring and ownership structure. Journal of Financial Economics, 49(1), 45–77.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Strategic Conception and Realization Path of the Construction of ' Western Hydrogen Energy Financial Trading Center ' in Chengdu-Chongqing Economic Zone

Pan Luo*, Jinglong Kang

Faculty of Law, University of Buenos Aires (UBA), Buenos Aires City, 1000, Argentina

Corresponding author: Pan Luo, luo533@est.derecho.uba.ar

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Against the dual backdrop of the deepening global energy transition and the rigid constraints of the "Dual Carbon" goals, hydrogen energy, as a clean, efficient, storable, and transportable secondary energy source, has become a core strategic pivot for ensuring national energy security and driving industrial upgrading. The Chengdu-Chongqing Economic Zone, leveraging its solid foundation for regional coordination, leading hydrogen energy industry layout, and policy dividends from building a Western Financial Center, possesses unique endowments and realistic conditions for jointly establishing a Western Hydrogen Energy Financial Trading Center. Based on theories of regional economic synergy, financial center formation, futures market functions, and electricity-hydrogen synergy, this paper systematically deconstructs the internal logic of deep integration between the hydrogen energy industry and financial markets, and comprehensively demonstrates the strategic value and feasibility of building such a center in the Chengdu-Chongqing Economic Zone. Research indicates that establishing a Western Hydrogen Energy Financial Trading Center can effectively address the developmental constraint of lacking a financial hub in the Chengdu-Chongqing region. By adopting a futures exchange model to build a trinity trading system encompassing "spot-forward-futures," it can foster a virtuous cycle of financial capital aggregation and hydrogen energy industry upgrading, truly realizing the core essence of finance serving the real economy. Furthermore, this paper proposes practical and forward-looking construction pathways from four dimensions: innovative trading product design, establishment of a multi-tiered market structure, construction of cross-regional coordination mechanisms, and improvement of a full-chain financial ecosystem. It aims to provide solid theoretical support and systematic practical guidance for the Chengdu-Chongqing region to become a national highland for hydrogen energy financial innovation and contribute to the new pattern of Western Development.

Keywords: Chengdu-Chongqing Economic Zone; Hydrogen Energy Industry; Financial Trading Center; Futures Market; Regional Synergy; Dual Carbon Goals

Published: Apr 30, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1326

1. Introduction

1.1 Research Background

1.1.1 The Prominent Strategic Position of Hydrogen Energy in the Global Energy Competition Landscape

In today's world, energy security has transcended mere supply assurance to become a key component of national core

competitiveness, with clean, low-carbon, and diversified development becoming an irreversible trend in global energy. The explosive growth of emerging industries like artificial intelligence and the digital economy has created exponential demand for computing power, the stability and sustainability of which fundamentally depend on the innovation and upgrading of the energy system. The latest "Global Carbon Budget" report released under the UNFCCC shows that the global carbon neutrality process has entered a critical phase, creating an urgent and pressing need to replace traditional fossil fuels. Hydrogen energy, as a secondary energy source with high energy density, zero-emission combustion, and diverse application scenarios, can not only effectively link the consumption and storage of intermittent renewable energy like wind and solar, solving structural challenges such as insufficient grid absorption capacity and cross-regional transmission bottlenecks, but also play an irreplaceable role in industrial decarbonization, heavy-duty transportation, and distributed energy storage. It has become a strategic high ground in global energy competition (Lei, 2025).

Against this backdrop, major global economies have incorporated the hydrogen energy industry into their national top-level strategies. The EU released its "Hydrogen Strategy", explicitly setting targets of 10 million tons of green hydrogen production capacity by 2030 and a pan-European hydrogen network by 2050. The US, through the "Infrastructure Investment and Jobs Act", allocated $7 billion specifically for building regional hydrogen hubs to construct a full hydrogen industry chain ecosystem. China, in its "14th Five-Year Plan for a Modern Energy System", clearly listed hydrogen energy as a strategic emerging industry, proposing to "build a hydrogen energy industry innovation ecosystem and promote the demonstration and application of hydrogen energy in transportation, industrial manufacturing, distributed energy, and other fields," aiming to build a full industry chain system covering "production-storage-transportation-utilization." The scaled development of the hydrogen energy industry is no longer merely a technological breakthrough issue but also requires comprehensive support from financial markets. As a core platform for resource allocation, the construction of a financial trading center will directly determine a country's pricing power and discourse power in the global hydrogen industry.

1.1.2 Concurrent Opportunities and Structural Bottlenecks in China's Hydrogen Energy Industry Development

Currently, China's hydrogen energy industry has transitioned from the technology R&D phase to the initial industrialization phase, making significant progress in production capacity scale, technological breakthroughs, and demonstration applications. By the end of 2024, China's annual hydrogen production reached 4.1 million tons, accounting for over 30% of global production, with industrial by-product hydrogen constituting about 80%. Green hydrogen capacity is expanding rapidly at an average annual growth rate exceeding 50%, and wind-solar hydrogen production demonstration projects have covered 12 provinces and regions including Inner Mongolia, Xinjiang, and Sichuan (Zhang & Lei, 2024). Regarding application scenarios, the demonstration and application of hydrogen energy in transportation are gradually scaling up. Over 350 hydrogen refueling stations have been built nationwide, and the number of hydrogen fuel cell vehicles in operation has exceeded 12,000, forming four major demonstration clusters in Beijing-Tianjin-Hebei, Yangtze River Delta, Pearl River Delta, and Chengdu-Chongqing.

However, the scaled development of China's hydrogen energy industry still faces three major structural bottlenecks. First, constraints in storage and transportation technologies: high-pressure gaseous storage and transportation entail high costs and energy consumption, liquid storage and transportation impose stringent requirements on technology and equipment, and new storage and transportation technologies like liquid organic hydrogen carriers and solid-state hydrogen storage are still in the early stages of industrialization, yet to achieve scaled application (Lei, 2025). Second, imperfect market mechanisms: hydrogen pricing lacks a market-based mechanism, with current prices heavily influenced by multiple factors such as industrial by-product hydrogen supply, policy subsidies, and transportation distance, leading to severe volatility and unstable expectations for upstream and downstream enterprises. Third, an inadequate financial support system: existing financial services mainly concentrate on traditional models like bank credit and government industrial funds, lacking specialized risk management tools such as futures and options. The capital aggregation effect has not yet formed, making it difficult to meet the high-investment, long-cycle, high-risk financing needs of the hydrogen energy industry. Particularly regarding financial trading platform construction, China has not yet established a national hydrogen trading market, hindering the industry's transition from demonstration applications to scaled development.

1.1.3 The Strategic Positioning and Developmental Imperatives of the Chengdu-Chongqing Economic Zone

The Chengdu-Chongqing Economic Circle, as a crucial strategic pivot in China's new "Dual Circulation" development pattern, serves as the core engine for the New Western Development Strategy in the new era and a vital growth pole for high-quality development. The "Chengdu-Chongqing Economic Circle Construction Plan Outline" explicitly states that "promoting the construction of the Chengdu-Chongqing Economic Circle is a major measure for building a new development pattern with domestic circulation as the mainstay and domestic and international circulations reinforcing each other." Jointly building a Western Financial Center is one of the core tasks in constructing the Chengdu-Chongqing Economic Circle (Liu, 2024). The Chengdu-Chongqing Economic Zone possesses unique endowment advantages in hydrogen energy industry development. Specifically: Firstly, Sichuan Province holds 23% of China's hydropower resources, 19% of wind energy resources, and 22% of solar energy resources, providing ample clean energy assurance for green hydrogen production. Secondly, Chongqing, as a traditional industrial base, produces over 300,000 tons of industrial by-product hydrogen annually with a purity exceeding 99.9%, providing low-cost hydrogen source support for the initial development of the hydrogen industry. Currently, the two regions have formed the embryo of a full hydrogen industry chain covering "production-storage-transportation-utilization" and are at the national forefront in areas such as wind-solar hydrogen production, fuel cells, and hydrogen refueling station construction.

However, compared to national financial centers like Beijing, Shanghai, and Shenzhen, the Chengdu-Chongqing region consistently lacks an influential financial trading hub. The existing financial market still relies heavily on bank credit and equity investment, with direct financing instruments like futures and bonds lagging. This limits the region's capacity to mobilize financial capital and the efficiency of resource allocation, failing to meet the demands of regional industrial upgrading and high-quality development (Feng et al., 2025). With the deepening implementation of the "Chengdu-Chongqing Economic Zone Regional Plan", cross-provincial cooperation mechanisms between Sichuan and Chongqing are continuously improving, effectively internalizing the spillover effects generated by regional competition, laying a solid institutional foundation for building a financial trading center (Feng et al., 2025). In this context, jointly building a Western Hydrogen Energy Financial Trading Center becomes an inevitable choice for the Chengdu-Chongqing region to overcome financial development bottlenecks, seize the strategic high ground in the hydrogen energy industry, and serve national energy security and the "Dual Carbon" goals.

1.2 Research Significance

1.2.1 Theoretical Significance

The theoretical contributions of this paper are mainly reflected in three aspects. First, it enriches the theoretical system of regional financial center construction by building a "industrial foundation-financial support-regional synergy-institutional guarantee" quadripartite theoretical framework for specialized financial trading centers. This breaks through the traditional focus on comprehensive financial functions in financial center research, providing a new analytical perspective for building cross-regional specialized financial trading platforms. Second, it expands the application boundaries of futures market functions by integrating the price discovery, risk hedging, and resource allocation functions of futures markets with the technical characteristics and development patterns of the hydrogen energy industry, enriching the theoretical connotation of new energy finance and providing theoretical support for the deep integration of emerging industries and financial markets. Third, it deepens the practical application of electricity-hydrogen synergy theory by exploring how financial markets can promote the scaled development of electricity-hydrogen synergy models through functions like capital allocation and risk mitigation, constructing a theoretical analytical framework for "technology-industry-finance" co-evolution.

1.2.2 Practical Significance

At the practical level, this research holds significant real-world value. First, it provides a feasible pathway for the Chengdu-Chongqing region to jointly build a Western Financial Center. Establishing a Hydrogen Energy Financial Trading Center would fill the gap in regional financial trading platforms, enhancing the status and influence of the Chengdu-Chongqing region in the national financial landscape. Second, it facilitates the high-quality development of the hydrogen energy industry. A market-based pricing mechanism would form hydrogen prices reflecting real supply-demand relationships,

providing specialized risk management tools for upstream and downstream enterprises to reduce operational risks, while simultaneously guiding social capital towards the hydrogen industry to solve financing difficulties. Third, it serves national strategy implementation by providing financial support for the new pattern of Western Development. Promoting the scaled development of the hydrogen energy industry aids in achieving the "Dual Carbon" goals and ensuring national energy security, fostering synergistic development between the regional economy and the ecological environment.

1.3 Research Content and Technical Route

1.3.1 Research Content

The core research content of this paper consists of three parts. First, systematically reviewing the development status of the hydrogen energy industry and financial trading markets, clarifying the technical characteristics, development trends, and financial demands of the hydrogen industry to build the theoretical foundation for the research. Second, comprehensively demonstrating the feasibility of building a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone from four dimensions: strategic significance, policy support, industrial foundation, and financial potential, clarifying its necessity and realistic conditions. Third, designing the construction framework for the Hydrogen Energy Financial Trading Center, including innovative trading product design, the establishment of a multi-tiered market structure, the construction of cross-regional coordination mechanisms, and the improvement of a full-chain financial ecosystem.

1.3.2 Technical Route

This paper adopts a logical approach of "theoretical review-feasibility analysis-pathway design." First, through the literature research method, it systematically reviews relevant theories such as regional economic synergy, financial center formation, futures market functions, and electricity-hydrogen synergy to construct the theoretical framework. Second, using the comparative research method, it compares experiences and lessons from domestic and international hydrogen energy financial markets and regional financial center construction to provide reference for the Chengdu-Chongqing region. Third, combining the actual situation of the Chengdu-Chongqing Economic Zone, it employs a combination of qualitative and quantitative analysis. By analyzing data such as regional hydrogen industry scale, total financial resources, and policy support intensity, it demonstrates the feasibility of construction and designs specific implementation pathways.

2. Literature Review

2.1 Research on Hydrogen Energy Industry Development

Foreign scholars started research on the hydrogen energy industry relatively early, focusing on technology R&D, industrialization pathways, and policy support. The IEA's (2023) "Hydrogen Outlook 2023" report indicates that hydrogen energy has irreplaceable advantages in industrial decarbonization, heavy-duty transport, and long-haul transportation, and its scaled development requires synergistic advancement of technological innovation, policy support, and infrastructure construction (IEA, 2023). Saeed et al. (2024), through empirical studies of major global hydrogen demonstration projects, show that the electricity-hydrogen synergy model can effectively increase the absorption rate of renewable energy and reduce the carbon emission intensity of the energy system, representing a core pathway to a zero-carbon energy system. Additionally, foreign scholars have conducted in-depth research on hydrogen storage and transportation technologies, fuel cell technologies, etc., providing theoretical support for technological breakthroughs in the hydrogen industry (Saeed et al., 2024).

Domestic scholars focus on industrial policy, regional layout, and technological innovation. Based on electricity-hydrogen synergy theory, Lei Xianzhang et al. (2025) propose the development principle of "prioritizing electricity, using electricity where possible, and hydrogen where suitable," considering liquid organic hydrogen carrier technology as a key path to solving hydrogen storage and transportation challenges, with the next 5 to 10 years being a critical phase for the maturation of the electricity-hydrogen synergy model (Lei, 2025). Zhang Changling and Lei Xianzhang (2024), from the dimensions of improving industrial policy and strengthening support measures, propose pathway suggestions for China's hydrogen industry development, emphasizing the important role of financial support in the industry's scaled development (Zhang & Lei, 2024). Research by Xie Feng et al. (2025) on methanol fuel cell technology indicates that the diverse application scenarios of hydrogen provide broad space for industry development, with different types of methanol fuel cells having their own

advantages in different application scenarios, suggesting optimization of technical routes based on application needs (Xie et al., 2025). Existing research has clarified the strategic value and development pathways of the hydrogen industry, but studies on the financial support system are relatively weak, especially lacking systematic exploration of hydrogen energy financial trading platform construction, failing to effectively bridge the technical demands of the hydrogen industry and the functional supply of financial markets.

2.2 Research on Regional Financial Center Construction

Research on regional financial center construction mainly revolves around formation conditions, development models, and economic effects. Kindleberger (1974), in "The Formation of Financial Centers: A Study in Comparative Economic History", proposed that the core function of a financial center is cross-temporal and spatial resource allocation, and its formation requires favorable location advantages, a strong industrial base, well-developed financial infrastructure, and a liberal institutional environment (Kindleberger, 1974). Porteous (1995), from the perspective of information asymmetry, argued that the formation of financial centers results from information aggregation and enhanced processing efficiency; the clustering of financial institutions can reduce information asymmetry, lower transaction costs, and improve resource allocation efficiency (Porteous, 1995).

Domestic scholars have conducted extensive empirical research on China's regional financial center construction. Liu Jun (2024) pointed out that jointly building a Western Financial Center in Chengdu-Chongqing is an important arrangement related to promoting regional coordinated development and building a new development pattern with domestic and international circulations reinforcing each other, and regional legislative synergy is key to breaking administrative barriers and realizing the free flow of financial resources (Liu, 2024). Feng Guoqiang et al. (2025), using the synthetic control method and taking the "Chengdu-Chongqing Economic Zone Regional Plan" as a quasi-natural experiment, empirically showed that cross-provincial regional development planning can effectively promote regional economic growth. The mechanism lies in mitigating the negative effects of regional competition through regional synergy and market integration, internalizing the spillover effects of regional competition (Feng et al., 2025). Huang Guanghong (2025) proposed that Chongqing should enhance the internationalization level of the Western Financial Center by playing the "international card" and strengthening connectivity with international financial markets. Existing research has confirmed the importance of regional synergy, industrial base, and institutional environment for financial center construction, but has not yet involved the construction of specialized financial trading centers in the new energy sector (Huang, 2025). Discussions on the feasibility, construction pathways, and coordination mechanisms for cross-regional joint construction of specialized financial trading platforms are insufficient.

2.3 Research on Futures Markets and New Energy Finance

Research on the application of futures markets in the new energy field mainly concentrates on carbon futures and new energy raw material futures. Deng Xiying (2025) designed contract terms for carbon emission allowance futures products using Hubei carbon quotas as the underlying asset, including core elements such as trading unit, delivery method, and quotation unit. The study argued that launching carbon futures could effectively expand carbon trading volume, enhance market liquidity, and through the price discovery function, provide strong support for China to gain pricing power in the international carbon trading market (Deng, 2025). Jiang Jianmei (2024), based on the actual situation of China's national carbon emission trading market, deeply explored carbon futures contract design and pricing research. Using the expectation theory pricing model combined with the ARIMA-GARCH model, a pricing formula for national carbon emission allowance futures products was derived, and a scientifically reasonable pricing scheme was obtained with the aid of Monte Carlo simulation techniques (Jiang, 2024).

Li Liu (2025), from the perspective of dynamic margins, designed carbon futures contracts suitable for the Chinese carbon market. Using the SPAN system to estimate market volatility and calculate Value at Risk (VaR), a dynamic margin mechanism was constructed, providing technical support for futures contract design (Li, 2025). Shang Wenqing (2024), through the case study of the Tsingshan nickel price surge incident, emphasized the importance of improving rules in the commodity futures market, providing references for optimizing trading mechanisms (Shang, 2024). Existing research has

accumulated experience in designing new energy futures products and building market structures, but research on building futures trading centers specifically for hydrogen energy—a unique energy product—remains a blank. It has failed to propose systematic construction plans by combining the technical characteristics of the hydrogen industry (e.g., difficult storage/transportation, high purity requirements) with the realistic needs of regional development.

2.4 Literature Review Summary

In summary, existing research has yielded rich results from the three dimensions of hydrogen energy industry, regional financial centers, and new energy futures, but there are still three main shortcomings: First, research perspectives are fragmented, lacking systematic studies that combine the technical characteristics of the hydrogen industry, the institutional advantages of regional synergy, and the functional positioning of financial trading platforms. An analytical framework for "technology-industry-finance-region" co-evolution has not been constructed. Second, research content tends to focus on a single field, with insufficient discussion on key issues such as the feasibility, construction pathways, and coordination mechanisms for cross-regional joint construction of specialized hydrogen energy financial trading centers, especially targeted research combining the actual conditions of specific regions (e.g., the Chengdu-Chongqing Economic Zone). Third, research methods lean towards theoretical analysis or single case studies, lacking comprehensive consideration of multiple factors such as regional industrial foundation, financial resources, and policy environment, resulting in proposed schemes with limited practicality.

This paper aims to address the above research gaps. Taking the Chengdu-Chongqing Economic Zone as the research object, it systematically analyzes the internal logic of deep integration between the hydrogen energy industry and financial markets, comprehensively demonstrates the strategic significance and practical feasibility of cross-regional joint construction of a hydrogen energy financial trading center, and designs practical and forward-looking construction pathways to provide references for related policy formulation and practical operations.

3. Theoretical Foundation

3.1 Regional Economic Synergy Theory

Regional economic synergy theory originates from the synergetics of systems science. Its core proposition is that a regional economic system is a complex system composed of multiple interrelated and interacting subsystems (e.g., industrial, financial, infrastructure subsystems). Through coordination, complementarity, and advantage-sharing among subsystems, the overall efficiency of the system can be maximized (Hagerstrand, 1968). This theory emphasizes that the key to regional synergy lies in breaking administrative barriers, promoting the free flow and optimal allocation of production factors (capital, technology, talent, data, etc.), and forming regional development synergy.

In China's regional development practice, cross-provincial economic zone construction is an important vehicle for achieving regional synergy. Empirical research by Feng Guoqiang et al. (2025) shows that the "Chengdu-Chongqing Economic Zone Regional Plan" effectively promoted economic growth in Sichuan and Chongqing through mechanisms such as market integration, resource sharing, and industrial collaboration, with a more pronounced promoting effect on Chongqing's economic growth than Sichuan's (Feng et al., 2025). This research confirms that cross-provincial cooperation can effectively internalize the spillover effects generated by regional competition, reduce competitive internal friction, and form synergistic development forces. Regional economic synergy theory provides important theoretical support for the joint construction of a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone: Sichuan and Chongqing have complementary strengths in hydrogen industry, financial resources, and policy environment. Through close cooperation in policy formulation, infrastructure construction, and regulatory coordination, regional financial resources and hydrogen industry resources can be effectively integrated, achieving a "1+1>2" synergistic effect and promoting the rapid establishment and healthy development of the Hydrogen Energy Financial Trading Center.

3.2 Financial Center Formation Theory

Financial center formation theory posits that the formation of a financial center results from the combined effects of multiple factors including location advantages, industrial base, financial infrastructure, and institutional environment (Kindleberger, 1974). Kindleberger proposed that financial centers have agglomeration and scale effects, attracting elements such as

financial institutions, talent, and capital, thereby reducing transaction costs, improving information transmission efficiency, and enhancing resource allocation efficiency. Porteous's (1995) information asymmetry theory further points out that the clustering of financial institutions can reduce information asymmetry, improve the efficiency and security of financial transactions, thereby attracting more financial activities to concentrate in the region (Porteous, 1995).

Combining China's financial center construction practice, the formation of a financial center requires three core conditions: First, strong industrial support, as industrial development generates substantial financial demand, providing abundant trading underlying assets and funding needs for financial markets. Second, well-developed financial infrastructure, including trading platforms, payment and clearing systems, credit reporting systems, etc., which form the foundation for normal financial market operations. Third, a favorable institutional environment, including policy support, regulatory systems, legal safeguards, etc., which can provide stable expectations for financial institution operations and financial transaction execution (Liu, 2024). The Chengdu-Chongqing Economic Zone, as the region with the largest economic aggregate and strongest industrial base in Western China, possesses the industrial support and location advantages for building a financial center. The construction of a Hydrogen Energy Financial Trading Center will fill the gap in financial infrastructure in the Chengdu-Chongqing region, improve the regional financial ecosystem, promote the aggregation of financial resources towards the hydrogen industry, and thereby enhance the region's position in the national financial landscape.

3.3 Futures Market Function Theory

The core functions of futures markets include price discovery, risk hedging, and resource allocation (Fama, 1970; Hull, 2018). The price discovery function refers to the formation of equilibrium prices reflecting real market supply and demand through open, fair, and efficient trading mechanisms where market participants trade based on their judgments of market supply-demand relationships, providing pricing references for the spot market (Fama, 1970). The risk hedging function allows market participants to transfer price fluctuation risks to speculators willing to bear them through hedging transactions, thereby stabilizing business expectations and reducing operational risks (Hull, 2018). The resource allocation function guides social capital towards advantaged industries and high-quality enterprises through price signals, optimizing resource allocation efficiency.

In the new energy field, the functions of futures markets are fully evident. Deng Xiying (2025) noted that launching carbon futures can effectively enhance carbon market liquidity, form price signals reflecting the external cost of carbon emissions through price discovery, and strengthen China's pricing power in international carbon trading (Deng, 2025). The institutional design and functional realization of futures markets provide a mature paradigm for building a Hydrogen Energy Financial Trading Center, offering the hydrogen industry a market-based pricing mechanism and resource allocation channel to promote its scaled development.

3.4 Electricity-Hydrogen Synergy Theory

Electricity-hydrogen synergy theory is a core theory for hydrogen energy industry development. Its central proposition is that by establishing a bidirectional coupling "electricity-hydrogen" energy utilization model, integrating renewable energy like wind and solar with hydrogen technology, a complete green energy ecosystem can be formed, enhancing the flexibility, stability, and sustainability of the energy system (Lei, 2025). This theory proposes the basic principle of "prioritizing electricity, using electricity where possible, and hydrogen where suitable," emphasizing the expansion of green hydrogen application scenarios and the R&D and application of liquid organic hydrogen carrier technology.

Electricity-hydrogen synergy theory provides technical pathway support for the scaled development of the hydrogen industry: On one hand, renewable energy power generation from sources like wind and solar is intermittent and fluctuating; direct grid integration can cause grid instability. Converting surplus electricity into hydrogen for storage through renewable energy-based hydrogen production can effectively solve grid absorption challenges. On the other hand, hydrogen storage, transportation, and application can extend the renewable energy industry chain, extending clean energy from the generation end to end-use energy consumption sectors like industry, transportation, and buildings, enhancing the decarbonization capability of the energy system (Lei, 2025). This theory indicates the technical feasibility of scaled hydrogen industry development. The construction of a financial trading center will provide funding support and market guarantees for promoting the electricity-

hydrogen synergy model, guiding capital through financial markets to promote the implementation of renewable energy hydrogen production projects and the industrial application of hydrogen storage and transportation technologies.

4. Feasibility Analysis of Jointly Building a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone

The construction of the western hydrogen energy financial trading center is not only of great significance to promote the sustainable economic development of Chengdu-Chongqing Economic Zone, but also of practical feasibility. This realistic feasibility is based on the rational judgment of the Chengdu-Chongqing Economic Zone at the four levels of strategy, policy, industry and finance.

4.1 Strategic Alignment: Resonating with National Energy Security and Regional Development Strategies

4.1.1 Seizing the High Ground in Global Hydrogen Industry Competition

As a core component of the future energy system, the construction of a financial trading center for hydrogen will directly determine a country's pricing power and discourse power in the global hydrogen industry. Currently, the global hydrogen market is still in its early development stage, lacking a unified pricing mechanism and trading platform, which provides an important window of opportunity for China to seize strategic initiative. Taking the lead in building a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone can attract global hydrogen industry resources (including technology, capital, talent, projects, etc.) to agglomerate in the region, forming a hydrogen pricing system and trading rules dominated by China, enhancing China's core competitiveness in global energy competition, and providing strong guarantees for national energy security.

4.1.2 Perfecting the Functional Layout of the Western Financial Center

The "Chengdu-Chongqing Economic Circle Construction Plan Outline"explicitly proposes to "jointly build a Western Financial Center," with financial trading centers being the core vehicle and important symbol of a financial center (Liu, 2024). Currently, China has formed three national financial centers in Beijing, Shanghai, and Shenzhen, with regional futures exchanges like Dalian, Zhengzhou, and Guangzhou having well-developed layouts, forming a nationwide financial trading network. However, the Chengdu-Chongqing region, as the area with the largest economic aggregate and strongest industrial base in Western China, consistently lacks an influential national financial trading platform. The existing financial market relies heavily on bank credit and equity investment, with direct financing instruments like futures and bonds lagging, severely limiting the region's capacity to mobilize financial capital and the efficiency of resource allocation (Feng et al., 2025). Building a Western Hydrogen Energy Financial Trading Center can fill this gap in the Chengdu-Chongqing region's financial trading hub, perfect the functional layout of the Western Financial Center, and enhance the region's status and influence in the national financial landscape.

4.1.3 Assisting in Achieving "Dual Carbon" Goals and Ensuring Energy Security

As a major global energy consumer and carbon emitter, China faces concurrent pressures of energy security and carbon reduction. The development of the hydrogen industry is a strategic pathway to ensure energy security and achieve the "Dual Carbon" goals (Zhang & Lei, 2024). The construction of a Western Hydrogen Energy Financial Trading Center can, through financial leverage, guide social capital to aggregate towards the hydrogen industry, promote the expansion of green hydrogen production capacity and the promotion of electricity-hydrogen synergy models, effectively reduce reliance on fossil fuels, and increase renewable energy absorption rates. Simultaneously, the scaled application of hydrogen can drive decarbonization in key sectors like industry and transportation, providing strong support for achieving national carbon reduction targets. Moreover, hydrogen resources in the Chengdu-Chongqing Economic Zone mainly come from renewables and industrial by-product hydrogen; developing the hydrogen industry can promote resource recycling, achieving synergistic progress in economic development and ecological protection.

4.2 Policy Support: National and Local Synergistic Efforts, Solid Institutional Guarantees

4.2.1 Continuous Release of National-Level Policy Dividends

In recent years, the state has intensively issued numerous policies supporting the hydrogen energy industry and regional

financial center construction, providing clear policy direction and institutional guarantees for building a Western Hydrogen Energy Financial Trading Center. Regarding the hydrogen industry, the "14th Five-Year Plan for a Modern Energy System" lists hydrogen energy as a strategic emerging industry, proposing to "build a hydrogen energy industry innovation ecosystem and promote the demonstration and application of hydrogen energy in transportation, industrial manufacturing, distributed energy, and other fields."The "Dual Carbon" Work Plan"explicitly proposes to orderly promote the development of the hydrogen energy industry and explore the application of green hydrogen in industry, transportation, and other fields." Regarding regional financial center construction, the "Chengdu-Chongqing Economic Circle Construction Plan Outline" explicitly supports jointly building a Western Financial Center in Chengdu-Chongqing, encouraging financial innovation and cross-regional synergy. The 2025 Central Economic Work Conference proposed to support the deep integration of new energy industries and financial markets, develop green finance, science and technology innovation finance, and promote industrial upgrading, providing policy basis for building a Hydrogen Energy Financial Trading Center.

4.2.2 Continuous Improvement of Local-Level Coordination Mechanisms

The governments of Sichuan and Chongqing attach great importance to the synergistic development of the hydrogen industry and finance, issuing a series of supporting policies and establishing sound coordination mechanisms. Sichuan Province issued the "Hydrogen Energy Industry Development Plan (2023-2027)", proposing to "build a hydrogen energy financial service system, support the establishment of hydrogen industry funds, and encourage financial institutions to innovate financial products and services." Chongqing released the "Western Financial Center Construction Action Plan (2023-2025)", listing "new energy financial innovation" as a key task, proposing to "support the construction of a hydrogen energy financial trading platform and promote the deep integration of the hydrogen industry and financial markets." Additionally, Sichuan and Chongqing have established the Chengdu-Chongqing Economic Circle Financial Synergistic Development Joint Conference System, making significant progress in policy coordination, market interconnection, regulatory coordination, and information sharing (Liu, 2024), providing a solid institutional foundation for jointly building the Hydrogen Energy Financial Trading Center.

4.3 Industrial Endowment: Embryonic Full Hydrogen Industry Chain Present, Strong Industrial Support

4.3.1 Significant Advantages in Energy Resource Endowments

The Chengdu-Chongqing Economic Zone possesses abundant renewable energy resources and industrial by-product hydrogen resources, providing ample hydrogen source guarantees for hydrogen industry development. Sichuan Province's theoretical hydropower potential reaches 143 million kW, ranking first nationally. Large hydropower stations like Baihetan and Wudongde have been built, providing ample clean energy for green hydrogen production. Wind energy resources are mainly distributed on the Western Sichuan Plateau, with developable capacity exceeding 30 million KW. Solar energy resources are relatively abundant in Western and Southern Sichuan, with annual irradiance reaching over 5000 MJ/m², possessing conditions for large-scale wind-solar hydrogen production. Chongqing, as a traditional industrial base, has well-developed steel, chemical, and other industries, producing over 300,000 tons of industrial by-product hydrogen annually with purity exceeding 99.9%, providing low-cost hydrogen source support for the initial development of the hydrogen industry.

4.3.2 Preliminary Formation of Full Industry Chain Layout

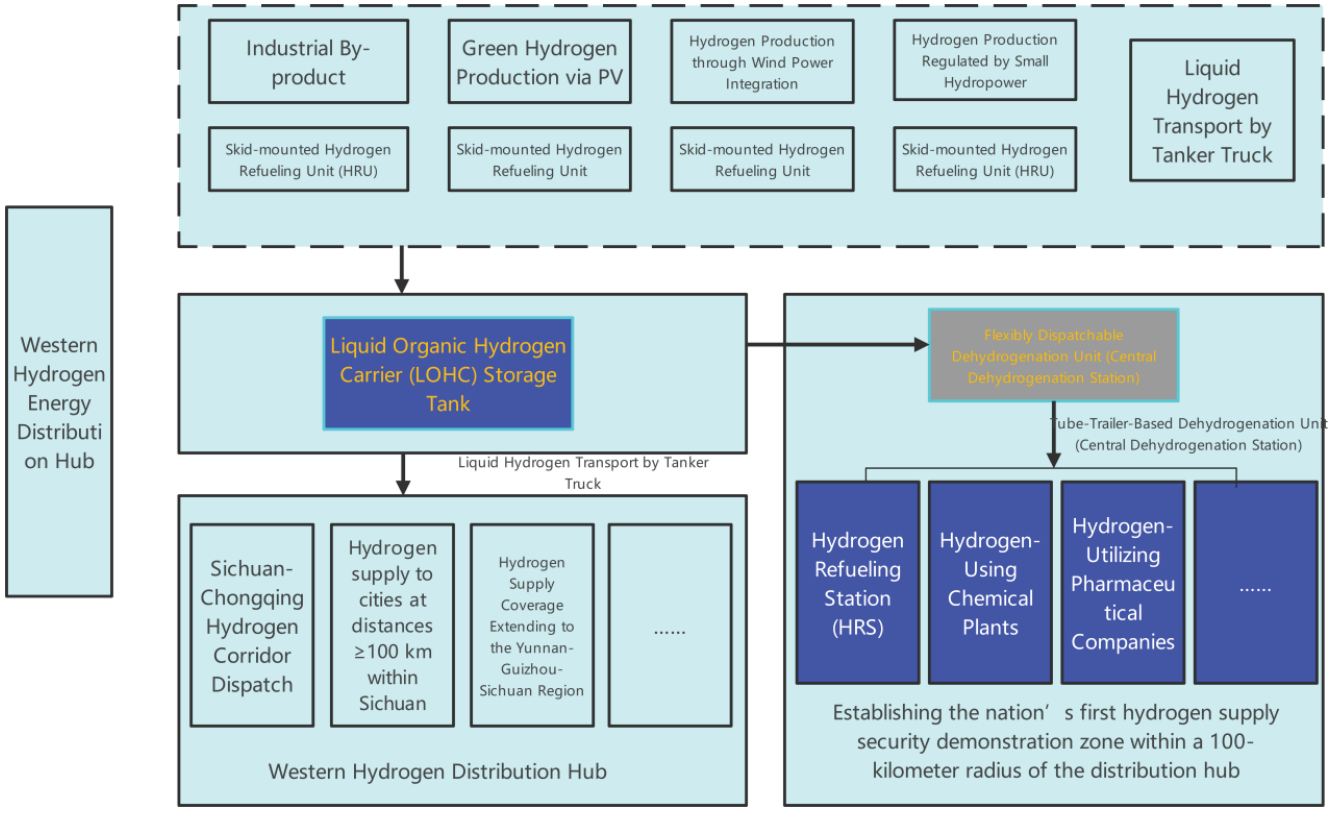
The Chengdu-Chongqing region has formed the embryo of a full hydrogen industry chain covering "production-storage-transportation-utilization," with a solid industrial foundation. In hydrogen production, enterprises like Sichuan Energy Investment and Chongqing Chemical & Pharmaceutical have established multiple wind-solar hydrogen production and industrial by-product hydrogen projects. Among them, the 500MW wind-solar hydrogen production project in Ganzi, Sichuan, is one of the largest green hydrogen production bases nationally, with an annual capacity of 20,000 tons. In storage and transportation, enterprises like Chongqing Qianwei Technology and Sichuan Academy of Aerospace Technology have made breakthroughs in high-pressure gaseous storage/transportation and liquid organic hydrogen carrier technologies, achieving scaled application. In fuel cells, enterprises like Chongqing Changan and Sichuan Hyundai have launched multiple hydrogen fuel cell vehicle models, with core components like fuel cell stacks and catalysts achieving localized production. In

applications, over 80 hydrogen refueling stations have been built in Chengdu and Chongqing, with the number of hydrogen fuel cell vehicles in operation exceeding 5,000, forming the nation's largest hydrogen demonstration application cluster.

4.3.3 Hydrogen Distribution Network Layout Aligns with Regional Development Needs

Lei Xianzhang proposed the concept of a "Western Hydrogen Distribution Center" in relevant reports, clarifying the overall layout of hydrogen supply, storage/transportation, and radiation, providing important guidance for cross-regional synergy in the Chengdu-Chongqing hydrogen industry. This concept covers core links including green hydrogen production (wind-solar hydrogen, industrial by-product hydrogen), storage/transportation systems (liquid truck transport, long-distance pipelines, hydrogen/oil storage tanks), radiation range (distant cities within Sichuan, Yunnan-Guizhou-Sichuan region), and application scenarios (hydrogenation industry, refueling stations, hydrogen-using enterprises), highly matching the hydrogen resource distribution and industrial layout of the Chengdu-Chongqing Economic Zone (see Figure 1).

Figure 1: Model Diagram of the Western Hydrogen Distribution Center (Lei, 2025)



This distribution network layout can effectively integrate hydrogen source resources and storage/transportation advantages in the Chengdu-Chongqing region, forming a closed-loop "production-storage/transportation-consumption" system. It provides a solid industrial vehicle for the Hydrogen Energy Financial Trading Center and lays the foundation for cross-regional hydrogen circulation and trading.

4.3.4 Strong Technological R&D Capabilities

The Chengdu-Chongqing region hosts numerous research institutions and universities, providing strong technological support for hydrogen industry development. Universities like Sichuan University and Chongqing University have established hydrogen energy research centers, undertaking multiple national key R&D programs in fuel cells, hydrogen production technologies, storage materials, etc. Research institutions like the Chengdu Institute of Organic Chemistry, Chinese Academy of Sciences, and Chongqing Materials Research Institute have achieved multiple patent results in key technologies like hydrogen storage materials and catalysts. The electricity-hydrogen synergy development pathway and Western Hydrogen Distribution Center concept proposed by Lei Xianzhang's team provide important technical guidance for the scaled development of the hydrogen industry (Lei Xianzhang, 2025). Additionally, the region has gathered a batch of innovative enterprises in the hydrogen field, forming a "university-research institution-enterprise" collaborative innovation technology

R&D system.

4.4 Financial Potential: Continuous Enhancement of Regional Financial Market Capacity, Sound Financial Ecosystem

4.4.1 Steady Growth in Total Financial Resources

In recent years, the financial market scale in the Chengdu-Chongqing region has continuously expanded, with total financial resources growing steadily, providing ample funding guarantees for building the Hydrogen Energy Financial Trading Center. By the end of 2024, the balance of domestic and foreign currency deposits of banking institutions in the region reached 25.6 trillion yuan, with the loan balance reaching 21.8 trillion yuan. Securities trading volume exceeded 120 trillion yuan, the number of equity investment institutions surpassed 1,500, with managed capital scale reaching 1.8 trillion yuan. The continuous growth of financial resources provides solid funding support for the construction and operation of the trading center and ample financing sources for hydrogen industry development.

4.4.2 Continuously Increasing Financial Innovation Vitality

The Chengdu-Chongqing region is at the national forefront in innovative practices in green finance, supply chain finance, sci-tech finance, etc., with continuously increasing financial innovation vitality. By the end of 2024, the balance of green credit in the region reached 3.2 trillion yuan, accounting for 14.7% of total loans. Green bond issuance reached 89 billion yuan, accounting for 8.3% of national issuance. National-level new areas like Chengdu Hi-tech Zone and Chongqing Liangjiang New Area have established hydrogen industry funds with total scale exceeding 20 billion yuan, providing full lifecycle financial support for hydrogen enterprises. Furthermore, financial institutions have innovatively launched financial products like hydrogen industry chain loans, financing leases, and intellectual property pledge loans, providing diversified financing channels for the hydrogen industry.

4.4.3 Increasingly Strong Market Demand

With the scaled development of the hydrogen industry, market demand for hydrogen financial services is becoming increasingly urgent. On one hand, hydrogen enterprises face high R&D investment and fixed asset investment, requiring diversified financing channel support. On the other hand, hydrogen prices are influenced by multiple factors like raw material prices, policy environment, and transportation distance, leading to severe fluctuations; upstream and downstream enterprises urgently need specialized risk management tools to hedge price risks. Estimates indicate that the hydrogen industry market scale in the Chengdu-Chongqing region reached 38 billion yuan in 2024 and is projected to exceed 200 billion yuan by 2030. The huge market scale provides solid demand support for building the trading center. Simultaneously, as a Western economic center, the Chengdu-Chongqing region's hydrogen market radiates to surrounding provinces like Yunnan, Guizhou, and Sichuan, providing broad market space for the trading center.

5. Construction Pathway for Jointly Building a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone

5.1 Trading Product Design: Building a Diversified and Differentiated Product System

5.1.1 Selection of Core Trading Underlying Assets

Drawing on design experience from new energy futures products like carbon emission allowance futures and lithium carbonate futures (Deng, 2025), and combining the technical characteristics and market demand of the hydrogen industry, the Western Hydrogen Energy Financial Trading Center should prioritize the selection of standardized, high-demand, significantly price-volatile hydrogen products as core trading underlying assets, specifically including:

1) Green Hydrogen: Using hydrogen produced from renewable energy (wind, solar, hydro, etc.) as the underlying asset, clarifying carbon footprint standards and certification systems for green hydrogen (e.g., lifecycle carbon intensity below 2 kg CO_2e/kg H_2) to meet green finance needs under "Dual Carbon" goals and attract environmentally-conscious enterprises and green investors.

2) Industrial By-product Hydrogen: Using hydrogen from steel, chemical, and other industries as the underlying asset, specifying quality standards like purity (e.g., above 99.9%) and impurity content, highlighting its cost advantage to serve initial market cultivation and meet demand for low-cost hydrogen sources from industrial users and the transportation sector.

3) Hydrogen Storage and Transportation Services: Using mainstream storage/transportation methods like liquid organic hydrogen carriers and high-pressure gaseous transport as underlying assets, standardizing elements like service price, transportation distance, and delivery time to address storage/transportation bottlenecks and promote cross-regional hydrogen circulation.

5.1.2 Futures Contract Terms Design

Futures contract terms should follow the principles of "standardization, operability, and risk controllability," fully considering the technical characteristics and market demand of the hydrogen industry. Specific term design is as follows:

1) Trading Unit: Set differentiated trading units based on application scenarios and market demand. For industrial users (e.g., chemical, steel enterprises), design large contracts of 10 tons/lot; for transportation sector users (e.g., hydrogen logistics companies, bus companies), design small contracts of 1 ton/lot to meet different user needs.

2) Delivery Method: Adopt a "physical delivery primary, cash delivery secondary" approach. Establish multiple delivery warehouses in the Chengdu-Chongqing region (e.g., in hydrogen industry clusters like Chengdu, Chongqing, Luzhou, Zigong), covering major production and consumption areas to facilitate physical delivery for enterprises. For trading varieties unsuitable for physical delivery (e.g., storage/transportation services), adopt cash delivery based on the delivery settlement price.

3) Quotation Unit: Use "yuan/kg" as the quotation unit. Referencing international hydrogen prices, domestic production costs, transportation costs, etc., set reasonable price fluctuation limits (e.g., daily price limits of ±5%) to balance market liquidity and trading stability.

4) Margin System: Introduce a dynamic margin mechanism (Li, 2025). Based on the SPAN system estimating market volatility and calculating Value at Risk (VaR), dynamically adjust margin ratios according to market fluctuations to balance capital usage and transaction security.

5) Delivery Months: Set monthly and quarterly delivery months. Monthly delivery meets short-term risk management needs; quarterly delivery meets long-term hedging needs, enhancing contract flexibility and applicability.

5.1.3 Derivative System Construction

Based on core futures products, gradually build a diversified derivative system to enrich trading varieties and meet different participant needs:

1) Hydrogen Options: Launch call and put options based on hydrogen futures, providing more flexible risk management tools for enterprises to hedge price risks according to their exposure.

2) Hydrogen Forward Contracts: Design customized forward contracts for large enterprises' long-term procurement needs, allowing parties to negotiate terms like price, delivery time, and location to lock in future hydrogen prices and stabilize business expectations.

3) Hydrogen ETFs: Issue Hydrogen Industry Exchange-Traded Funds (ETFs) tracking a hydrogen industry index, attracting ordinary investors to participate in hydrogen industry investment and enhancing market liquidity.

4) Carbon-Hydrogen Linked Products: Integrate with the carbon emission trading market to design linked financial products, e.g., "carbon allowance+green hydrogen" portfolio products, encouraging enterprises to reduce emissions and increase green hydrogen use for synergistic progress in carbon reduction and hydrogen industry development.

5.2 Market Structure Establishment: Building a Multi-tiered Market System with "Spot-Forward-Futures" Linkage

5.2.1 Market Tier Design

Drawing on construction experience from domestic mature futures exchanges like the Guangzhou Futures Exchange (Yang, 2024), build a multi-tiered hydrogen financial trading market system with a trinity of "spot market - forward market - futures market," where each tier complements and interacts with others:

1) Spot Market: Positioned as an immediate delivery market, providing standardized spot hydrogen trading services to meet short-term procurement needs. Its core function is forming market-based spot prices, providing a pricing basis for forward and futures markets. The trading center should establish a well-developed spot trading platform for electronic trading,

settlement, and delivery to reduce costs.

2) Forward Market: Positioned as a medium-to-short-term risk management market, providing 1-12 month forward trading services for enterprises. Its core function is helping enterprises lock in future prices and stabilize expectations. The center should formulate standardized forward contract templates while allowing parties some flexibility to negotiate terms, balancing standardization and customization.

3) Futures Market: Positioned as a long-term pricing and risk management market, launching standardized futures contracts. Its core functions are price discovery and risk hedging. Through open, fair, and efficient trading mechanisms, it forms equilibrium prices reflecting long-term supply-demand relationships, guiding social capital towards the hydrogen industry.

Linkage among the three tiers is achieved through arbitrage mechanisms: When spot and futures prices deviate, arbitrageurs conduct arbitrage trading between markets, bringing prices back to reasonable ranges and ensuring effective pricing. Simultaneously, the forward market acts as a bridge, smoothing price fluctuations and enhancing market stability and liquidity.

5.2.2 Participant Cultivation

Diversified participants are key to market activity. The trading center should attract various market entities including industry players, financial institutions, investors, and government/regulators to form a multi-tiered, diversified participation pattern:

1) Industry Players: Include hydrogen producers, storage/transportation companies, and end-users. Producers can lock in selling prices via futures to hedge against price declines; end-users can lock in purchase prices to hedge against increases; storage/transport firms can hedge transportation cost fluctuation risks.

2) Financial Institutions: Include commercial banks, securities companies, futures companies, insurers, etc. Banks can provide financing and settlement services; securities/futures firms can offer brokerage and risk management services; insurers can develop hydrogen-related insurance products (e.g., storage/transport liability, equipment damage insurance) to provide risk coverage.

3) Investors: Include public/private funds, Qualified Foreign Institutional Investors (QFII), and individual investors. Their participation enhances liquidity and aids price discovery. The center should strengthen investor education for rational investment.

4) Government & Regulators: Government departments encourage participation through policy guidance and funding support; regulators maintain order and ensure standardized operation through rule-making and enforcement.

5.2.3 Trading Mechanism Optimization

Adopt advanced trading mechanisms to enhance market efficiency and transparency, ensuring fairness, impartiality, and efficiency:

1) Trading Mode: Use electronic auction trading for fully electronic trading, reducing costs and improving efficiency. The platform must have robust technical support to handle large-scale concurrent trading with stability and security.

2) Matching Mechanism: Adopt a “price-time priority” matching mechanism to ensure fairness. Among orders meeting transaction conditions, the best-priced order executes first; for same price, the earliest-submitted order executes first.

3) Information Disclosure: Establish a well-developed information disclosure system specifying content, methods, timing, etc. The center should timely publish market info like production, inventory, demand, prices, as well as participant positions and transaction data, reducing information asymmetry and enhancing transparency.

4) Clearing & Settlement: Establish an independent clearinghouse using netting clearing to reduce counterparty risk. It should have a sound risk reserve fund system to cover losses in case of participant default, ensuring normal market operation.

5.3 Regional Coordination Mechanism: Building a Cross-Regional “Policy Synergy-Infrastructure Co-construction-Regulatory Linkage” System

5.3.1 Policy Synergy Mechanism

Drawing on experience from Chengdu-Chongqing Economic Circle financial synergy development (Liu, 2024), establish a cross-regional policy synergy mechanism to break administrative barriers and form policy synergy:

1) Establish a Joint Working Group: Composed of financial authorities from both sides (e.g., PBC Chengdu Branch,

Chongqing Office; CBIRC Sichuan & Chongqing bureaus; CSRC Sichuan & Chongqing bureaus), along with development & reform, energy, industry & information technology departments, to coordinate and advance the center's construction/operation and resolve major issues.

2) Unify Policy Standards: Formulate unified standards for market access, tax incentives, funding support, talent policies, etc., to avoid fragmentation. E.g., provide tax reductions for participating hydrogen enterprises, rental subsidies for settled institutions, and preferential settlement, housing policies for hydrogen finance professionals.

3) Share Policy Resources: Integrate policy resources like industrial funds and fiscal subsidies from both sides to jointly establish a Hydrogen Financial Innovation Fund supporting product innovation and market cultivation. Also, create a policy information sharing platform for timely release of local policies.

5.3.2 Infrastructure Co-construction

Strengthen co-construction and sharing of hydrogen industry and financial infrastructure in the Chengdu-Chongqing region to reduce costs and improve operational efficiency:

1) Co-build Hydrogen Storage/Transport Network: Jointly build cross-regional hydrogen pipeline networks and liquid organic hydrogen carrier storage bases, deeply aligning with the Western Hydrogen Distribution Center concept (Lei, 2025), forming a supply network covering Chengdu-Chongqing and radiating to Yunnan-Guizhou-Sichuan. Unify technical standards and safety norms for storage/transport to ensure smooth and safe cross-regional flow.

2) Share Financial Infrastructure: Share payment/clearing systems, credit reporting systems, financial regulatory platforms, etc., to achieve real-time fund clearing, interconnection of participant credit information, and sharing of regulatory data, lowering transaction and regulatory costs.

3) Delivery Warehouse Layout: Establish delivery warehouses in hydrogen industry clusters like Chengdu, Chongqing, Luzhou, Zigong, forming a region-wide delivery network. Warehouses should comply with national standards, have well-developed storage facilities, safety systems, and quality inspection equipment to ensure delivered commodity quality and safety.

5.3.3 Regulatory Linkage Mechanism

Establish a cross-regional regulatory linkage mechanism to maintain market order and ensure healthy development:

1) Unify Regulatory Standards: Follow the principle of "same market, same rules, same regulation" to formulate unified standards/rules, clarifying regulatory responsibilities and processes. Standards should cover market access, trading conduct, information disclosure, etc., ensuring comprehensiveness and effectiveness.

2) Information Sharing Platform: Establish a regulatory information sharing platform for real-time sharing and interconnection of regulatory data. Regulators can query participant transaction data, positions, fund flows, etc., to promptly identify potential issues.

3) Joint Enforcement Mechanism: Form joint enforcement teams for collaborative investigation and penalty of violations. Upon discovering market manipulation, insider trading, misrepresentation, etc., regulators from both sides should closely cooperate in investigation and penalties to maintain order.

4) Risk Warning Mechanism: Jointly establish market risk warning models using big data, AI, etc., to monitor price fluctuations, trading volume changes, fund flows in real-time, promptly identify and address potential problems. Issue risk warnings and take necessary measures to stabilize the market during anomalies.

5.4 Financial Ecosystem Improvement: Building a Full-Chain "Financing Support-Service Provision" Financial Ecosystem

5.4.1 Financing Support System

Provide full lifecycle financing support for the hydrogen industry to solve financing difficulties and promote upgrading:

1) Equity Financing: Encourage hydrogen enterprises to list on capital markets like Sci-Tech Innovation Board, ChiNext, Beijing Stock Exchange. Support cooperation between the trading center and these exchanges to provide listing guidance, roadshow services, etc. Also, support PE/VC funds to invest in hydrogen enterprises for equity financing.

2) Debt Financing: Support hydrogen enterprises in issuing green bonds, supply chain finance bonds, medium-term notes, etc.

Encourage banks to offer specialized loans with preferential rates and flexible terms. Encourage leasing companies to provide equipment financing lease services to ease fixed asset investment pressure.

3) Industrial Funds: Expand the scale of hydrogen industry funds in Chengdu-Chongqing, establish a Hydrogen Financial Innovation Fund focusing on supporting technology R&D, storage/transport infrastructure, and demonstration projects. Adopt a "government-guided, societal participation" model to attract social capital.

4) Policy Financial Support: Leverage policy banks to provide long-term, low-cost funding for the hydrogen industry. E.g., China Development Bank, Export-Import Bank of China can offer specialized loans for infrastructure and cross-border hydrogen trade.

5.4.2 Service Provision System

Improve financial service support to enhance market service capabilities and ensure healthy development:

1) Intermediary Services: Attract intermediary agencies like accounting/law firms, asset appraisers, credit rating agencies, quality inspection agencies to provide professional services. E.g., audit, legal consultation/contract review, quality inspection for delivered commodities.

2) Research & Consulting: Establish a Hydrogen Finance Research Center to conduct research on the industry and financial markets, providing decision-making consulting for participants. Regularly release industry development reports, market analyses, policy interpretations, etc., for business and investment reference.

3) Talent Cultivation: Collaborate with universities like Sichuan University, Chongqing University, Southwestern University of Finance and Economics to establish hydrogen finance programs, cultivating composite talent with both industry and finance knowledge. Provide on-the-job training and continuing education to enhance the professional quality of existing practitioners.

4) International Cooperation: Strengthen cooperation with international bodies like the International Hydrogen Energy Association and foreign futures exchanges to introduce advanced experience and technology, enhancing the center's internationalization. Promote cross-border hydrogen trade and financial transactions to attract foreign enterprises and investors.

6. Conclusion and Outlook

6.1 Research Conclusion

Based on regional economic synergy theory, financial center formation theory, futures market function theory, and electricity-hydrogen synergy theory, this paper systematically demonstrated the strategic significance and practical feasibility of jointly building a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone, proposing a construction pathway characterized by "product diversification, market multi-tiering, all-round synergy, and full-chain ecosystem." The main conclusions are as follows:

Jointly building a Western Hydrogen Energy Financial Trading Center in the Chengdu-Chongqing Economic Zone holds significant strategic importance. First, it aligns with multiple national strategic needs including energy security, "Dual Carbon" goals, and Western Development. Its construction can help China seize the global hydrogen industry competitive high ground, perfect the Western Financial Center's functional layout, solve the development bottleneck of lacking a financial hub in Chengdu-Chongqing, and provide financial support for high-quality hydrogen industry development. Second, the Chengdu-Chongqing Economic Zone possesses a solid foundation for building the center. Policy support at national and local levels provides institutional guarantees; a leading hydrogen industrial base provides industrial support; a growing financial market provides financial support. Third, the center should be constructed following a futures exchange model, building a multi-tiered "spot-forward-futures" linked market system, designing a diversified, differentiated trading product and derivative system, establishing a cross-regional "policy synergy-infrastructure co-construction-regulatory linkage" coordination mechanism, and improving a full-chain "financing support-service provision" financial ecosystem to ensure its healthy operation and functional realization.

6.2 Research Outlook

This research provides theoretical support and practical guidance for jointly building the center, but some limitations remain.

Future research can deepen in the following aspects:

Regarding quantitative research, this paper primarily uses qualitative analysis. Future work could construct econometric models to quantitatively analyze the center's driving effects on regional economic growth, hydrogen industry development, employment, etc., and factors affecting market pricing efficiency. Regarding product pricing, this paper proposed pricing ideas. Future research could combine the industry's cost structure and market supply-demand relationship to build more precise pricing models for scientific basis.

In summary, building the Western Hydrogen Energy Financial Trading Center is a systematic project requiring joint efforts from national policy support, local government coordination, industry participation, financial institution cooperation, and research institution support. Promoting its construction will provide powerful financial support for China's hydrogen industry high-quality development and energy transition, making significant contributions to achieving the "Dual Carbon" goals and national energy security.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Lei, X. (2025). Electricity-hydrogen synergy: Opening a new chapter for the high-quality development of wind and solar new energy. Natural Gas and Oil, 43(05), 1–7+175.

[2] Zhang, C., & Lei, X. (2024). Ideas and pathway suggestions for promoting the development of China's hydrogen energy industry. China Development Observation, (11), 123–128.

[3] Liu, J. (2024). On regional legislative coordination in the joint construction of the Western Financial Center by Chengdu and Chongqing. Economic Law Forum, 33(02), 170–183.

[4] Feng, G., Yan, Z., & Ma, T. (2025). The economic impact of the Chengdu-Chongqing Economic Zone on Sichuan and Chongqing—Also on the implications for forming a new pattern in the new era of Western Development. Journal of Chongqing University (Social Science Edition), 31(05), 93–105.

[5] IEA. (2023). Hydrogen Outlook 2023. Paris: International Energy Agency.

[6] Saeed, M., Abbas, A., Ali, M., et al. (2024). Electro-hydrogen synergy: A pathway to high renewable energy penetration. Renewable and Sustainable Energy Reviews, 168, 112897.

[7] Xie, F., Meng, H., Lei, X., Li, X., Ding, F., & Shao, Z. (2025). Research progress on high-efficiency methanol fuel cell technology. Energy Storage Science and Technology, 14(11), 4237–4244. https://doi.org/10.19799/j.cnki.2095-4239.2025.0314.

[8] Kindleberger, C. P. (1974). The Formation of Financial Centers: A Study in Comparative Economic History. Princeton: Princeton University Press.

[9] Porteous, D. J. (1995). The Geography of Finance: Spatial Dimensions of Intermediary Behaviour. Aldershot: Edward Elgar Publishing.

[10] Huang, G. (2025, February 19). Chongqing plays the "international card" for the Western Financial Center. Chongqing Daily, 004. https://doi.org/10.28120/n.cnki.ncqrb.2025.000722.

[11] Deng, X. (2025). Product design of Hubei carbon emission allowance futures (Master's thesis, Guangdong University of Technology). https://doi.org/10.27029/d.cnki.ggdgu.2025.003697.

[12] Jiang, J. (2024). Product design of national carbon emission allowance futures (Master's thesis, Guangdong University of Foreign Studies). https://doi.org/10.27032/d.cnki.ggdwu.2024.001060.

[13] Li, L. (2025). Design of China's carbon futures from the perspective of dynamic margins(Master's thesis, Shanghai Normal University). https://doi.org/10.27312/d.cnki.gshsu.2025.002151.

[14] Shang, W. (2024). Research on risk control in the commodity futures market for users(Master's thesis, Guangxi

University). https://doi.org/10.27034/d.cnki.ggxiu.2024.000069

[15] Li, B., & Cui, Z. (2025). Construction of the Chengdu-Chongqing Economic Zone, Chengdu-Chongqing Urban Agglomeration, and Chengdu-Chongqing Economic Circle.China Investment, (Z7), 78–80.

[16] Hagerstrand, T. (1968). Diffusion of Innovations as a Spatial Process. Chicago: University of Chicago Press.

[17] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383–417.

[18] Hull, J. C. (2018). Options, Futures, and Other Derivatives (11th ed.). Upper Saddle River: Prentice Hall.

[19] Yang, Y. (2024, December 11). Building a world-class futures exchange to better serve economic and social development. Financial Times, (007).

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

The Impact of Green Finance on Regional Total Factor Productivity

Qisong Wang[1], Shangjun Li[2], Yajie Han[3]*

1. Zhuhai College of Science and Technology, Zhuhai, Guangdong, 519041, China
2. College of Business, City University of Hong Kong, Hong Kong, China
3. Zhuhai Fudan Innovation Institute, Hengqin, Guangdong, 519031, China

Corresponding author: Yajie Han, yajiehan@fudan-zhuhai.org.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Guided by the dual carbon goals of peaking emissions and achieving carbon neutrality, Guangdong Province, as the vanguard of China's reform and opening-up, offers valuable green finance development experiences that are of reference significance for South China and nationwide. Utilizing panel data from 21 prefecture-level cities in Guangdong (2000-2019), this study employs the entropy method and stochastic frontier analysis (SFA) to calculate a city-level Green Finance Development Index (GF) and Total Factor Productivity Change (TFPC). A two-way fixed effects panel model is then applied to empirically analyze the impact mechanisms of green finance on regional productivity. By decomposing total factor productivity into technological progress (TC), technical efficiency change (TEC), scale efficiency change (SC), and allocative efficiency change (AEC), the research delves into the operational pathways of green finance. Results show its complex, multidimensional effects: while it limits technical and scale efficiency, it significantly enhances technological progress and resource allocation efficiency.

Keywords: Green Finance; Total Factor Productivity; Stochastic Frontier Analysis; Sustainable Development
Published: May 7, 2026
DOI: https://doi.org/10.62177/apemr.v3i3.1323

1. Introduction

With rapid economic development, environmental degradation, ecological deterioration, and resource scarcity have become increasingly prominent, making green development a critical agenda in national development strategies. During the National Ecological Conservation Conference in July 2023, China's head of state emphasized the need to refine economic policies supporting green and low-carbon development through enhanced fiscal, tax, financial, and pricing measures to accelerate the society-wide green transition. In recent years, the Chinese government has progressively enhanced its green finance policy system, utilizing novel instruments such as carbon emission reduction facilities to channel financial resources into clean energy and energy efficiency sectors, thereby creating new drivers for high-quality economic growth. China's economic growth model has transitioned from the extensive expansion during the planned economy era, to factor-driven growth following reform and opening-up, and subsequently to innovation-driven growth in the new era. This transition has seen total factor productivity improvements become the fundamental force propelling high-quality economic growth. As a crucial element of China's economic competitive advantage, green finance now acts as a key catalyst for driving the progress of contemporary national economic development.

A pioneer in China's economic reform, Guangdong has balanced growth with ecological goals, emerging as an innovator in green finance. In 2017, the PBOC Guangzhou Branch launched the "Single Hub, Dual Spheres" framework, centering on the Guangzhou Green Finance Innovation Pilot Zone. This model leverages Greater Bay Area connectivity and spatial linkages to spur provincial green finance development. Coupled with the 2023 Carbon Peak Plan's strategic measures, green finance is now central to Guangdong's low-carbon transition. Through sustained supply-side reforms over the past decade, the province has made significant advances in green finance innovation. As a key tool for sustainable development, green finance promotes high-quality growth and shows strong potential to boost regional total factor productivity (TFP), a core efficiency metric that informs policy and resource optimization. The critical role of green finance in the Bay Area's carbon neutrality efforts underscores the importance of studying Guangdong's mechanisms and their TFP effects, an inquiry with considerable academic and policy value.

Against this backdrop, this study empirically analyzes the impact of green finance development on total factor productivity (TFP) and its underlying mechanisms, using panel data from 21 prefecture-level cities in Guangdong Province (2000–2019). Our primary objectives and key contributions are threefold. First, by deconstructing the green finance system into seven key components, we systematically examine their potential effects on TFP. This analysis provides both theoretical support and empirical evidence for the effective implementation of regional green finance policies. Second, moving beyond the conventional three-component decomposition of TFP, which comprises technological progress, technical efficiency, and scale efficiency, this paper introduces allocative efficiency as a fourth dimension. It further explores how green finance development affects local resource allocation efficiency, thereby establishing a more comprehensive, four-dimensional analytical framework. Third, we elucidate the pathway mechanisms through which green finance influences productivity in Guangdong. These findings offer substantial financial policy insights for advancing the Green and Beautiful Guangdong Initiative and the national “Dual Carbon” strategy, thereby contributing to sustainable green finance development.

2. Theoretical Analysis and Hypotheses Development

Financial development is closely tied to economic growth, with green finance increasingly becoming a crucial driver for sustainable development. Decision-makers and policymakers in developing countries need to acknowledge the importance of financial development and its role in the economic growth process, as it impacts employment, population, inequality, and poverty (Chavriya et al., 2024). While focusing on development, constructing a well-rounded financial development system can reduce ecological footprints, with this effect being particularly evident in both developing and developed nations (Aslam et al., 2023). Along with a solid financial development framework, financial innovation and the quality of financial institutions are key factors influencing economic development. Strong financial innovation promotes financial growth, while lower institutional quality in emerging markets can hinder this development (Alawi et al., 2022). The rise of green finance has brought new energy to the socio-economy, offering fresh perspectives for individuals, businesses, financial institutions, and policymakers. As early as the 1990s, environmental protection and social responsibility began to gain attention in foreign businesses and the financial sector. People started realizing the impact of financial activities on the environment and society, and the concept of green finance gradually evolved and matured. The development of green finance has become the first step in enhancing sustainability in capital markets, changing the interaction patterns among market participants (Maltais & Nykvist, 2021). These shifts in interaction have impacted the sustainability of internal business operations. This influence, however, extends beyond individual companies and affects economies in a multidimensional manner. Falcone (2020) suggests that the development of green finance not only accelerates the transition of economies to sustainable development but also creates favorable conditions to ensure a fair competitive environment for both traditional and green economies. Furthermore, any support from green finance tools is likely to raise regional GDP per capita, both in the short and long term (Rasoulinezhad & Taghizadeh-Hesary, 2022).

It is easy to see that the development of green finance has some positive impacts on economic growth, including attracting investment resources, enhancing regional economic competitiveness, optimizing industrial structure transformation, promoting ecological improvement, creating new jobs, and boosting international competitiveness (Wei, 2024). First, green finance directly influences the movement of market funds and resources. Within the context of green finance, financial

institutions help improve economic quality by optimizing the direction of credit allocation. Relevant policies guide financial capital into environmental protection and green construction, promoting urban transformation and upgrading. De Haas & Popov (2023) studied how the financial structures of different countries affect the green transformation of industries. They found that the stock market promoted green innovation in carbon-intensive sectors, reducing carbon emissions per unit of output and reallocating investments to more energy-efficient industries, underscoring the critical role of capital markets in promoting green finance and resource allocation. Thus, advancing green finance requires leveraging capital markets and capital's dominant role in resource allocation. Capital's leading role drives resources toward green development, channeling more funds into environmental industries and indirectly contributing to green economic development. This process, involving the allocation of green finance funds and industrial upgrades, directly fuels high-quality green economic growth and regional ecological improvement (Yang, 2023).

Moreover, green finance mobilizes savings, allowing financial institutions to attract more deposits, ensuring capital formation, and optimizing factor allocation across industries. This increases capital efficiency, enabling enterprises to invest more in technology adoption and transformation, thus raising regional innovation and management levels (Palmaccio et al., 2023). This indirectly reflects that green finance policies can directly influence businesses. The implementation of these policies increases capital accessibility for enterprises, lowers financing costs, and facilitates technological progress through adjusted innovation input structures. Luo & Wang (2024) found that green finance development lowered the capital threshold for the commercialization of research outcomes in intermediate product enterprises, raising demand for patent technology. At the same time, the commercialization cycle of research results shortened, enhancing research efficiency and output in R&D departments. Zhang et al. (2024) discovered a positive correlation between corporate green innovation and skill premiums. Their analysis emphasized the significant role of green finance policies in increasing the use of skilled labor, especially in highly urbanized areas. As companies use more capital, they increase demand for skilled labor, which raises skill premiums, attracting labor from surrounding areas to regions with more advanced financial development, thus influencing changes in regional labor factors.

On the flip side, the rapid development of green finance may also bring some potential drawbacks to economic growth. Because the development of green finance requires significant investment and innovation, society faces economic fluctuations while the demand for patented technology and different labor factors in regions continues to increase, which may exacerbate the shortage and competition for funds and talents, thereby hindering the green transition of traditional enterprises. He et al. (2022) found that in the process of promoting the green transition of enterprises, the effect of green finance on optimizing the economic structure is not as significant as expected. The rapid advancement of policies has instead imposed very high transition costs on traditional enterprises that support the national economy and weakened confidence in the green transition. The inflow of green funds is often biased, and traditional enterprises and non-energy-saving environmental protection enterprises usually find it difficult to obtain sufficient financial support, thus facing external financing constraints (Matteo et al., 2023). This increases their R&D costs, causing them to gradually lose confidence in the transition, negatively affecting the overall economic growth of the region. For different types of enterprises, the vigorous development of green finance may also have varying degrees of negative impact. Especially for high-pollution enterprises, the rapid advancement of green finance-related policies has subjected them to credit constraints and increased their exit risk, while also inhibiting their technological innovation (Zhang et al., 2024). Even though the relevant policies have improved the investment efficiency of some enterprises, they have suppressed the performance of others, indirectly inhibiting the total factor productivity of enterprises. Therefore, the development of green finance will not only positively impact economic development but may also present obstacles. Thus, we propose:

Hypothesis 1a (H1a): The development of green finance in Guangdong Province has a positive impact on regional total factor productivity;

Hypothesis 1b (H1b): The development of green finance in Guangdong Province negatively impacts regional total factor productivity.

In recent years, various studies have emerged exploring the relationship between green finance and economic growth, with

varying conclusions. Therefore, the impact of green finance on total factor productivity growth is likely to be integrated and complex, necessitating a clear understanding of its mechanisms and their heterogeneity. Based on the work of Kumbhakar & Lovell (2000), this paper decomposes total factor productivity into four components: technological progress, technical efficiency, scale efficiency, and allocative efficiency. The paper examines the impact of green finance on each of these components and constructs a theoretical framework for analyzing the relationship between green finance and regional total factor productivity from an empirical perspective.

First, technological progress plays a crucial role in sustainable economic development and is a key factor in reducing energy consumption in businesses. It relies on the efficient allocation of funds and innovation in technology research and development. Technological progress is not only essential for improving energy efficiency across industries but also a significant determinant of socio-economic growth, interacting with energy consumption and human capital (Oliveira et al., 2024). The development of green finance allows enterprises to access funds more easily and at lower costs, enabling them to invest more in structural adjustments, which leads to changes in regional technological progress. With the support of green funds and policies, enterprises are encouraged to research, introduce, and apply environmentally friendly and sustainable technologies, which in turn promotes regional technological innovation and progress (Wu et al., 2022). Additionally, green finance supports the improvement of services and products offered by financial institutions, fostering innovation and enabling the creation of more diverse green financial products. High-quality and dynamic financial institutions are crucial for ensuring financial stability and long-term sustainable development, as well as aligning financial development with economic policies for coordinated growth (Sajid et al., 2023). Given this, does the support from green funds and policies have a positive impact on technological progress in Guangdong, which is at the forefront of China's green finance development? According to Zeng et al. (2023), the development of green finance significantly promotes technological innovation in the environmental protection industry, reducing capital barriers and fostering technological progress. Lin, Lam et al. (2023) found that the establishment of green finance reform and innovation pilot zones has a lagged effect on promoting sustainable economic and social development. The positive impact on sustainable development is significant only in regions with lower financial development, lower industrialization, and higher technological innovation levels. On the other hand, compared to traditional financial systems, green finance systems are stricter in financing high-pollution and high-energy-consuming industries, leading to capital outflow from these sectors. Guo et al. (2024) found that financial institutions continuously raise the loan thresholds for "three high" enterprises, limiting their interest-bearing debt financing and new investments. They also set capital market indicators to guide social funds away from "three high" enterprises and toward green enterprises, further shrinking the space for high-energy-consuming, high-pollution, and high-emission enterprises. While allocating funds in the market, traditional industries receive less support compared to energy-saving and environmentally friendly industries. The costs and risks of their R&D and innovation technologies increase, so green finance may lead to capital outflows and risk aversion, thus hindering technological progress and suppressing total factor productivity. Thus, we propose:

Hypothesis 2a (H2a): The development of green finance in Guangdong Province has a positive impact on regional technological progress;

Hypothesis 2b (H2b): The development of green finance in Guangdong Province has a negative impact on regional technological progress.

Technical efficiency reflects the ability to convert inputs into effective outputs and is an important factor influencing total factor productivity. When green finance development makes it easier for enterprises to access financial support, it will inevitably affect the adjustment of their input structures and strategies, thereby impacting output and reflecting technical efficiency. The inflow of green funds alleviates the financing constraints faced by green enterprises, providing strong support for technological innovation and transformation in the green sector, as well as the introduction of advanced management models, which continuously improve technical efficiency and promote the commercialization of research outcomes, enhancing overall regional R&D efficiency (Lei et al., 2024). However, small and medium-sized enterprises (SMEs) that are not energy-efficient or environmentally friendly will be forced to transform but may not obtain sufficient financial support, resulting in high risks and uncertain returns in their innovation activities. This makes them more susceptible to external

financing constraints, which limits their technological R&D efficiency in the green sector and hinders regional technical efficiency progress (Rupeika-Apoga & Petrovska, 2022). Thus, we propose:

Hypothesis 3a (H3a): The development of green finance in Guangdong Province has a positive impact on regional technical efficiency;

Hypothesis 3b (H3b): The development of green finance in Guangdong Province has a negative impact on regional technical efficiency.

The increase in capital investment is an important manifestation of the expansion of economic and production scales. The flow and allocation of funds brought about by the development of green finance undoubtedly have an impact on regional scale efficiency. Firstly, the improvement in the allocation efficiency of green credit funds across different industries has promoted the upgrading of urban industrial structures. Green credit improves the allocation efficiency of funds in the renewable energy industry, promoting its rapid development. At the same time, it produces a crowding-out effect on other traditional industries, urging them to accelerate their transformation and upgrading to meet the requirements of green development (Wang & Fan, 2023). Second, green innovative products and services satisfy consumers' diverse and multi-layered needs for sustainability, health, safety, and environmental protection, such as energy-efficient appliances, organic food, and new energy vehicles (Li & Zhang, 2023). Consumers are increasingly concerned not only with the basic functions and quality of products but also with their environmental impact and sustainability. These products and services have not only accelerated the optimization and upgrading of urban industrial structures but have also maximized scale effects and structural dividends, injecting new vitality into industrial development and improving regional total factor productivity. However, for enterprises of different scales, China’s financial system has traditionally favored serving state-owned and large enterprises with lower production efficiency, some of which even face over-service. The support from green funds may cause large-scale state-owned and large enterprises to become even more bloated, deviating further from the optimal enterprise size, which could reduce regional scale efficiency (Wang & Shao, 2024). Therefore, the development of green finance does not necessarily lead to the full realization of scale efficiency. Thus, we propose:

Hypothesis 4a (H4a): The development of green finance in Guangdong Province has a positive impact on regional scale efficiency.

Hypothesis 4b (H4b): The development of green finance in Guangdong Province has a negative impact on regional scale efficiency.

The development of green finance has driven the flow and further allocation of funds and resources, while allocation efficiency reflects the effectiveness of how funds and resources are distributed within the economy. Green finance policies guide funds from high-pollution, high-energy consumption, and overcapacity industries toward green industries, while also directing capital from inefficient enterprises to more efficient ones within these high-pollution sectors. This drives the upgrading of industrial structures and urban transformation (Gu et al., 2021). As society shifts toward emerging economic models, it may also impact the resource allocation within traditional enterprises. Funds and resources tend to favor energy-saving and environmentally friendly enterprises, which can increase the cost of converting resources into output for traditional enterprises, even those with strong growth potential and development prospects but without an energy-efficient or environmentally friendly focus (Hu et al., 2024). The rapid development of green finance increases the transformation costs for traditional enterprises that support the national economy, leading to a gradual decline in their confidence in transitioning to green development. This, in turn, negatively impacts the efficiency of resource allocation within these enterprises, making both the enterprises and the region gradually deviate from the optimal allocation efficiency. Thus, we propose:

Hypothesis 5a (H5a): The development of green finance in Guangdong Province has a positive impact on regional allocation efficiency;

Hypothesis 5b (H5b): The development of green finance in Guangdong Province has a negative impact on regional allocation efficiency.

3. Study Design

3.1 Model Construction

First, a baseline econometric model is established to analyze the impact of green finance development in Guangdong Province on total factor productivity (TFP). Following the Hausman test, a fixed effects model is selected. Considering the heterogeneity across years and regions, a two-way fixed effects model is constructed as follows:

$$TFPC_{i,t} = \beta_0 + \beta_1 GF_{i,t} + \sum \beta_j Control_{it} + \gamma_t + \mu_i + \varepsilon_{i,t} \quad (1)$$

In Equation (1), the dependent variable is $TFPC_{i,t}$, the independent variable is $GF_{i,t}$, and $Control_{it}$ represents a series of control variables. γ_t denotes time fixed effects, μ_i denotes regional fixed effects for each prefecture-level city, and $\varepsilon_{i,t}$ is the random error term.

3.2 Variable Selection and Explanation

3.2.1 Sample Selection and Data Sources

Since the outbreak of COVID-19, the real economy has been severely impacted, reducing the effective labor force and indirectly disrupting the normal market transaction mechanism. The pandemic has disrupted the market economy, severely squeezing both supply and demand, which in turn affects the flow and allocation of capital in the market. The pandemic reduced the effective labor force and indirectly disrupted the normal market transaction mechanism. For the total factor productivity studied in this paper, both capital and labor, as important influencing factors, are inevitably affected by the pandemic. In addition, in terms of finance, Goldstein, Koijen & Mueller (2021) discovered that the pandemic led to an increase in unemployment rates, which triggered higher default rates in lending, thereby increasing financial risks and seriously impacting financial institutions and the macroeconomy. Moreover, the uncertainty caused by the pandemic also affected investor sentiment and led to changes in corporate stock prices (Kirci Altinkeski, Cevik & Dibooglu, 2022; Al-Yahya, 2022). Therefore, to ensure the reference value of the research results, this study excludes the years of the pandemic and post-pandemic period, focusing on the period from 2000 to 2019 for the 21 prefecture-level cities in Guangdong Province, to explore whether the promotion of green finance development in Guangdong has any adverse impact on economic development. Data sources include the China Science and Technology Statistical Yearbook, China Energy Statistical Yearbook, China Financial Yearbook, China Industrial Statistical Yearbook, China Tertiary Industry Statistical Yearbook, statistical yearbooks from cities in Guangdong Province, environmental bulletins from each city, national economic and social development statistical bulletins, and authoritative websites of institutions such as the People's Bank of China.

3.2.2 Dependent Variables

The dependent variable is the growth rate of total factor productivity (TFPC) for 21 prefecture-level cities in Guangdong Province. Stochastic Frontier Analysis (SFA) is adopted to construct the model, as it allows comprehensive and stochastic sample selection while decomposing TFP into multiple components. The basic functional form is:

$$Y = f(X)e^{v-u} \quad (2)$$

Here, Y represents actual output, measured by "gross regional product" with 2000 as the base year. X denotes factor inputs (e.g., labor and capital). Capital (K_t) is calculated using the perpetual inventory method:

$$K_t = K_{t-1} \times (1 - 0.096) + m$$

where K_t is the capital stock in period t, K_0 is the base-year capital stock (2000), m represents fixed asset investment at constant prices, and the depreciation rate adopts 9.6% from Zhang, Wu & Zhang (2004). Labor (L) is measured by total employment, calculated by summing the number of urban unit employees and private individual employees. v is the random error term reflecting systemic inefficiency, and u is the technical inefficiency term. Following Battese & Coelli (1992), a translog stochastic frontier production function is specified:

$$\begin{aligned} \ln y_{it} = \ln f(K,L,t) = \beta_0 + \beta_K \ln K_{it} + \beta_L \ln L_{it} + \beta_t t + 0.5\beta_{KK}(\ln K_{it})^2 + 0.5\beta_{LL}(\ln L_{it})^2 \\ + \beta_{KL}(\ln K_{it})(\ln L_{it}) + \beta_{Kt}(\ln K_{it})t + \beta_{Lt}(\ln L_{it})t + \beta_{tt}t^2 + v_{it} - u_{it} \end{aligned} \quad (3)$$

Here, Y denotes economic output, K is capital, L is labor, subscripts i and t represent region and year, and β are estimated parameters. Following Kumbhakar & Lovell (2000), TFPC is decomposed into:

Technological Change (TC)

Technical Efficiency Change (TEC)

Scale Efficiency Change (SC)

Allocative Efficiency Change (AEC)

The aggregate output growth rate is expressed as:

$$GY = GA + GK + GL \tag{4}$$

where *GY* is the economic growth rate, *GA* is the TFP growth rate (TFPC), and *GK* and *GL* are the growth rates of capital and labor, respectively. TFPC is indirectly derived as:

$$GA = GY - GK - GL \tag{5}$$

Taking the derivative with respect to time *t* (subscripts omitted for simplicity):

$$TFPC = \frac{dlnY}{dY} \cdot \frac{dY}{dt} - \sum_j (\lambda_j \cdot \frac{dX_j}{dt}) \tag{6}$$

Expanding Equation (3) yields:

$$\frac{dlnY}{dY} \cdot \frac{dY}{dt} = \frac{\partial lnf(X,t)}{\partial t} + \sum_j \frac{\partial lnf(X,t)}{\partial (\ln X_j)} \cdot \frac{d(\ln X_j)}{dt} - \frac{\partial U}{\partial t} \tag{7}$$

$$\frac{dlnY}{dY} \cdot \frac{dY}{dt} - \sum_j \lambda_j \cdot \frac{d(\ln X_j)}{dt} = \frac{\partial lnY}{\partial t} - \frac{\partial U}{\partial t} + \sum_j \varepsilon_j \cdot \frac{d(\ln X_j)}{dt} - \sum_j \lambda_j \cdot \frac{d(\ln X_j)}{dt} \tag{8}$$

$$TFPC = \frac{\partial lnY}{\partial t} - \frac{\partial U}{\partial t} + \sum_j \varepsilon_j \cdot \dot{X}_j - \sum_j \frac{\varepsilon_j}{RTS} \cdot \dot{X}_j \tag{9}$$

$$TFPC = \frac{\partial lnY}{\partial t} - \frac{\partial U}{\partial t} + (RTS - 1) \sum_j \lambda_j \cdot \dot{X}_j \tag{10}$$

In Equation (10):

$$\lambda_j = \frac{\varepsilon_j}{RTS} \; ; \; \varepsilon_j = \frac{\partial \ln f(x_j)}{\partial (\ln x_j)} \; ; \; \dot{X}_j = \frac{d(\ln X_j)}{dt}$$

Here:

- $\frac{\partial lnY}{\partial t}$ represents TC (technological change).
- $\frac{\partial U}{\partial t}$ represents TEC (technical efficiency change).
- $(RTS - 1) \sum_j \lambda_j \cdot \dot{X}_j$ represents SC (scale efficiency change).

Thus, TFPC can be decomposed as follows:

(1)Technological Change (TC): TC is defined as the rate of change in the technological frontier after controlling for factor inputs. The formula is:

$$TC_{it} = \frac{\partial lnf(X,t)}{\partial t} = \beta_t + \beta_{Kt}(\ln K_{it}) + \beta_{Lt}(\ln L_{it}) + 2\beta_{tt}t \tag{11}$$

(2)Technical Efficiency Change (TEC): Following Leibenstein (1966), TEC is defined as the rate of change in production inefficiency over time. The formulas are:

$$TEC_{it} = \frac{\partial lnTE_{it}}{\partial t} = \frac{TE_{it} - TE_{it-1}}{TE_{it-1}} \tag{12}$$

$$TE_{it} = \frac{\partial U}{\partial t} = \exp(-u_{it}) \tag{13}$$

The technical inefficiency term u_{it} in Equation (13) is derived from the estimation of Equation (3).

(3)Scale Efficiency Change (SC): SC is defined as the rate of change in production efficiency due to scale effects. The formula is:

$$SC_{it} = (RTS - 1) \sum_j \lambda_j \dot{X}_j = (RTS_{it} - 1)(\lambda_{Kit} x_{Kit} + \lambda_{Lit} x_{Lit}) \tag{14}$$

Here:

RTS denotes returns to scale.

λ_j represents the output elasticity of input j relative to overall returns to scale.

ε_j represents the output elasticity of input j on the production frontier.

$\dot{X}_j$ denotes the growth rate of factor inputs, including x_K (capital) and x_L (labor).

(4)Allocative Efficiency Change (AEC): AEC is defined as the rate of change in resource allocation efficiency over time. The formula is:

$$AEC_{it} = \sum_j (\lambda_j - s_j)\dot{X}_j = (\lambda_{Kit} - s_{Kit})x_{Kit} + (\lambda_{Lit} - s_{Lit})x_{Lit} \tag{15}$$

Allocative efficiency is measured by the deviation between λ_j (output elasticity of input j) and s_j (cost share of input j). Optimal allocation occurs when $\lambda_j = s_j$. Here, s_K is the cost share of capital, and s_L is the cost share of labor, with s_K+s_L=1.

The total factor productivity change (TFPC) is calculated as:

$$TFPC = TEC + TC + SC + AEC \quad (16)$$

Using Stata, TFPC and its components are estimated for Guangdong Province. Descriptive statistics are shown in Table 1.

Table 1. Descriptive Statistics of TFPC and Its Decomposition in Guangdong Province

Name	Sample Size	Min	Max	Mean	Std. Dev.	Median
SC	399	-3.326	0.552	-0.265	0.313	-0.228
TC	399	-0.099	0.143	0.004	0.040	0.003
TEC	399	-0.018	0.201	0.017	0.029	0.007
AEC	399	-0.186	0.384	-0.030	0.050	-0.027
TFPC	399	-2.969	0.762	-0.275	0.303	-0.259

Based on the calculation results, further exploration is conducted on the structure and structural changes of TFPC and its components in Guangdong Province, as shown in Figure 1.

Figure 1: Temporal trends of TFPC and its components in Guangdong province (2001–2019)



According to the dynamic changes of various decomposition indicators over time shown in Figure 1, the growth dynamics of TFPC and its components in Guangdong Province from 2001 to 2019 generally show a decreasing trend, indicating a loss of economic growth efficiency in recent years.

The technological progress (TC) was positive during the period from 2001 to 2011 but showed a slow downward trend. After 2012, it became negative, indicating that Guangdong's technological frontier output has been gradually declining, and the benefits from investments in technological innovation are diminishing. Technological efficiency improvement (TEC) has been positive, showing that recent improvements in technology efficiency have promoted TFPC. However, it has maintained a steady downward trend, approaching zero. Resource allocation efficiency improvement (AEC) also shows a downward trend, remaining negative since 2005. The declines in TEC and AEC suggest that recent factor inputs in Guangdong have not been converted into effective output. However, after 2015, although still negative, there was slight improvement, likely due to the increased influx of capital, talent, and high-tech, making the resource inputs more effective over time. The scale efficiency improvement (SC) has been negative for many years, indicating that changes in scale efficiency hinder TFPC, primarily due to environmental and error-related factors. The process can be divided into four stages: the first stage, from 2001 to 2004, saw a continuous decline in scale efficiency, possibly due to the extensive development problems arising as Guangdong's economy shifted from slow growth in the 1990s to rapid growth, with increasing economic openness and technological improvements in eastern and western Guangdong. The second stage, from 2004 to 2008, showed slight improvement, likely because Guangdong focused on supporting impoverished and mountainous areas and investing in infrastructure. The third stage,

from 2008 to 2011, was impacted by the global financial crisis, causing a sharp decline followed by recovery, which affected scale efficiency. The fourth stage, from 2011 to 2019, saw a U-shaped trend in scale efficiency, first declining and then increasing. This may have been caused by increased investments in economic revitalization in Guangdong's eastern, western, and northern regions, leading to extensive development and a decrease in scale efficiency. However, the subsequent rapid growth in the number of legal entities, employees, and corporate assets, along with government investments in infrastructure, industrial park construction, urban expansion, and the large-scale application of high-tech, led to improvements in scale efficiency.

3.2.3 Independent Variable

This paper selects the green finance development level (GF) of 21 prefecture-level cities in Guangdong Province as the independent variable. The entropy method is employed, and Excel is used for raw data processing, measurement, and comprehensive evaluation. With the rising prominence of sustainable development topics, related literature has proliferated, and the evaluation system for green finance development levels has diversified. To ensure the accuracy and scientific rigor of the measurement results, this study prioritizes representative and quantifiable indicators for GF measurement. The specific indicators are listed in Table 2.

Table 2. Green finance development level measurement indicator system

Primary Indicator	Secondary Indicator	Indicator Nature	Calculation Method	Weight Coefficient
Green Finance Development Index	Green Credit	Positive	Environmental Project Credit Amount / Total Credit Amount	9.48%
	Green Investment	Positive	Environmental Pollution Control Investment / GDP	22.69%
	Green Insurance	Positive	Environmental Liability Insurance Revenue / Total Premium Income	9.45%
	Green Bonds	Positive	Green Bond Issuance Volume / Total Bond Issuance Volume	9.59%
	Green Support	Positive	Fiscal Environmental Protection Expenditure / General Fiscal Budget Expenditure	26.95%
	Green Funds	Positive	Green Fund Market Value / Total Fund Market Value	10.22%
	Green Equity	Positive	Carbon, Energy Use Rights, and Pollution Rights Trading / Total Equity Market Trading Volume	11.63%

The evaluation methods and indicator properties in the table above reference the works of Zhang et al. (2024). For green credit, some scholars use green credit balances or credit balances for high-energy-consuming and polluting industries (Hou et al., 2023). However, considering data availability and the representative nature of environmental project credit ratios, this paper adopts the ratio of environmental project credit to total credit. Green investment directly addresses environmental degradation and reflects regional efforts to promote green finance; thus, it is measured by the ratio of environmental pollution control investment to GDP. Green insurance is measured by the share of environmental liability insurance revenue in total premium income. Green bonds reflect capital market support for ecological development and are measured by the ratio of green bond issuance to total bond issuance. Green support, indicating governmental fiscal commitment to environmental protection, is measured by the ratio of fiscal environmental protection expenditure to general budget expenditure. Green funds and green equity are measured by their respective market shares, capturing financial market engagement in green initiatives.

To eliminate dimensional differences and standardize the data, dimensionless processing is applied before entropy method

calculations. The steps are as follows:

(1)Comprehensive Evaluation Formula:

$$S_j = \sum_{j=1}^{m} w_j X_{ij}^* \tag{17}$$

Where, S_j represents the comprehensive evaluation value; w_j ($0 \le w_j \le 1$) represents the weight of each indicator; X_ij^* represents the evaluation object after dimensionless processing; i (i = 1, …, n) represents the i-th evaluated object, and j (j = 1, …, m) represents the value of the j-th indicator of the evaluated object.

(2)Dimensionless processing is applied to all the positive indicators in the indicator system, and the data is normalized to standardize all indicators, resulting in X_{ij}^*. The normalization process is as follows:

$$X_{ij}^* = \frac{X_{ij} - \min X_{ij}}{\max X_{ij} - \min X_{ij}} \tag{18}$$

(3)Translation processing is applied to ensure the internal patterns of the data remain intact after standardization, by adding 0.0001 to each value:

$$y_{ij} = X_{ij}^* + 0.0001 \tag{19}$$

(4)The proportion of the j-th indicator value of the i-th evaluated object to the total sum of the indicators for that object is calculated as follows:

$$P_{ij} = \frac{y_{ij}}{\sum_{j=1}^{m} y_{ij}} \tag{20}$$

(5)The entropy value for each indicator is calculated using the following formula:

$$P_{ij} = \frac{y_{ij}}{\sum_{j=1}^{m} y_{ij}} \tag{21}$$

Where $k = \frac{1}{\ln n}$, $k > 0$, and $0 \le e_j \le 1$.

(6)The information utility value, also known as the information divergence coefficient, is calculated to represent the data's information and volatility. The formula is:

$$d_j = 1 - e_j \tag{22}$$

The weight of each indicator is calculated using the formula:

$$w_j = \frac{d_j}{\sum_{j=1}^{m} d_j} \tag{23}$$

Finally, the comprehensive score for each evaluated object is calculated using:

$$S_j = \sum_{i=1}^{m} (w_j \cdot X_{ij}^*) \tag{24}$$

Based on these calculations, the weight calculation results for each evaluated object in GF are summarized in Table 3.

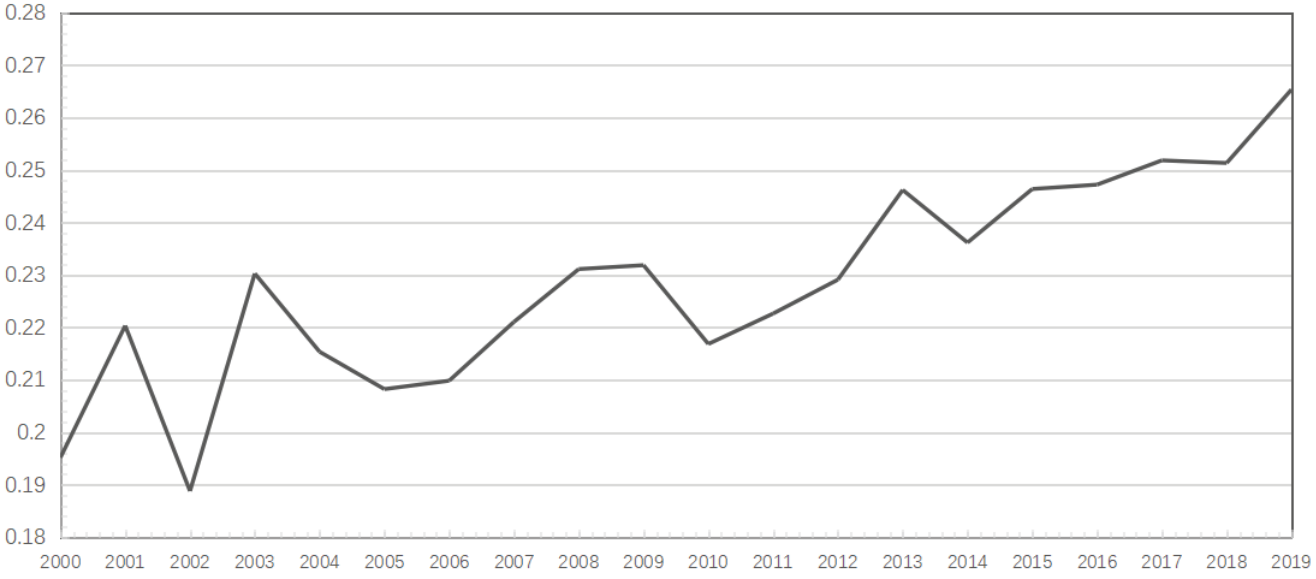
Table 3. The weight calculation results for various indicators in GF:

Indicator	Entropy	Utility Value	Entropy Weight
Green Credit	0.9793	0.0207	9.48%
Green Investment	0.9505	0.0495	22.69%
Green Insurance	0.9794	0.0206	9.45%
Green Bonds	0.9791	0.0209	9.59%
Green Support	0.9413	0.0587	26.95%
Green Funds	0.9777	0.0223	10.22%
Green Equity	0.9747	0.0253	11.63%

The results obtained using the entropy method show that Green Investment and Green Support have weights of 22.69% and 26.95%, respectively, indicating that these two are the main factors influencing the level of green financial development. Both indicators reflect the government's financial investment in environmental pollution control and protection, which directly promotes green development. Their significant share in the comprehensive evaluation indicates that the primary driving force for Guangdong's green financial development from 2000 to 2019 came from government policies, demonstrating the province's commitment to implementing green development plans and its ongoing exploration of new green development models and pathways.

Based on the calculated results, further investigation into the dynamic changes of the GF comprehensive score over time is shown in Figure 2.

Figure 2. Dynamic temporal changes in the green finance development level in Guangdong province (2001–2019)



Based on the dynamic changes of GF over time shown in Figure 2, from 2000 to 2019, Guangdong's GF generally showed an increasing trend, with a slight decline in 2002 followed by a recovery in 2003. Afterward, there was a downward trend for two years. After 2005, it increased slowly in five-year cycles and slightly declined afterward. However, the overall trend still showed growth, indicating that the green financial development in Guangdong Province has achieved significant results under the implementation of green development plans in various prefecture-level cities across the province.

3.3 Control Variables

The data range for the control variables in this study is from 2000 to 2019 for the 21 prefecture-level cities in Guangdong Province. To control for the effects of other factors on regional total factor productivity, this study follows the approach of Liu & Zhu (2024), ultimately adopting the following indicators: (1) Economic Development: The logarithm of the per capita GDP adjusted for inflation is used as a measure, where the ratio of the actual GDP in 2000 to the total population at the end of the year serves to control the nonlinear impact of economic development on total factor productivity; (2) Government Intervention: Measured by the ratio of fiscal expenditure to GDP; (3) Environmental Regulation: Measured by the ratio of total investment in environmental pollution control to GDP; (4) Urban Scale: Measured by taking the logarithm of the registered urban population; (5) Unemployment Rate: Measured by the urban registered unemployment rate; (6) Industrial Structure: Measured by the proportion of the value added by the secondary and tertiary industries to GDP; (7) Education Expenditure Level: Measured by the ratio of education expenditure to the local government's general public budget expenditure. The raw data for each variable come from city statistical yearbooks, publications by statistical bureaus, work reports of various cities, and the national economic and social development statistical bulletins of each city. For a small amount of missing data, linear interpolation was used to fill the gaps.

Table 4. Definition of Key Variables and Descriptive Statistics

Indicator	Variable Name	Variable Abbreviation	Sample Size	Mean	Standard Deviation	Min	Max
Dependent Variables	Total Factor Productivity	TFPC	399	-0.275	0.303	-2.969	0.762
	Technological Progress	TC	399	0.004	0.04	-0.099	0.143
	Technological Efficiency Progress	TEC	399	0.017	0.029	-0.018	0.201
	Scale Efficiency Progress	SC	399	-0.265	0.313	-3.326	0.552
	Allocation Efficiency Progress	AEC	399	-0.03	0.05	-0.186	0.384
Independent Variable	Green Finance Development Level	GF	399	0.23	0.07	0.047	0.444
Control Variables	Economic Development	pgdp	399	6.585	9.105	0.3	45.621
	Government Intervention	gov	399	0.129	0.063	0.048	0.399
	Environmental Regulation	env	399	0.035	0.029	0.002	0.239
	Urban Scale	scale	399	5.875	0.529	4.33	6.861
	Unemployment Rate	ump	399	2.573	0.423	0.5	3.76
	Industrial Structure	isT	399	0.884	0.088	0.59	1.068
	Education Expenditure Level	eel	399	0.21	0.053	0.014	0.477

4. Empirical Results and Analysis

4.1 Baseline Regression Results

As shown in Table 5, control variables are sequentially added from columns (1) to (8), ultimately establishing the baseline model estimation results. The parameter estimates reveal that after controlling all variables, the GF's coefficient on TFPC is statistically significant and negative at the 1% confidence level, indicating that green finance development negatively impacts total factor productivity, which aligns with hypothesis H1b's theoretical predictions.

Table 5. Baseline regression of green finance development on total factor productivity

Variable	TFPC							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GF	0.025	-0.063	-0.185***	-0.187***	-0.181***	-0.172***	-0.173***	-0.167***
	(0.45)	(-1.14)	(-2.79)	(-2.82)	(-2.81)	(-2.67)	(-2.70)	(-2.60)
pgdp		0.784***	0.706***	0.870***	1.413***	1.268***	1.226***	1.202***
		(5.30)	(4.77)	(5.20)	(6.96)	(5.83)	(5.65)	(5.54)
env			0.354***	0.366***	0.365***	0.366***	0.411***	0.386***
			(3.20)	(3.32)	(3.41)	(3.42)	(3.80)	(3.55)
gov				0.366***	0.365***	0.366***	0.411***	0.386***
				(3.32)	(3.41)	(3.42)	(3.80)	(3.55)

Variable	TFPC							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
scale					-1.571***	-1.497***	-1.463***	-1.490***
					(-4.48)	(-4.25)	(-4.18)	(-4.27)
ump						0.138*	0.134*	0.156**
						(1.83)	(1.79)	(2.07)
isT							0.199**	0.232**
							(2.18)	(2.50)
eel								0.133*
								(1.86)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.135	0.198	0.220	0.229	0.271	0.277	0.287	0.294
Sample size	399	399	399	399	399	399	399	399

Note: *, **, and *** indicate that the estimated coefficients are significant at the 10%, 5%, and 1% significance levels, respectively; same table below.

Combining the results of the simple regression estimation, it can be seen that GF does not show a facilitating effect on TFPC, but instead has a certain degree of inhibiting effect. For this reason, the causes of this problem will be further explored in this paper.

4.2 Mechanism Tests

In order to better explore the mechanism of action of the negative effect of GF on TFPC, TFPC was decomposed and regression estimation was performed for each component separately to analyze the mechanism of the effect of GF on TC, TEC, SC, and AEC, respectively, and the results are shown in Table 6.

Table 6. Regression estimation results of green financial development on each decomposition component of total factor productivity

Variable	TC		TEC		SC		AEC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GF	0.242***	0.108***	-0.092**	-0.142***	-0.027	-0.196***	0.175***	0.207***
	(8.43)	(4.59)	(-2.14)	(-2.80)	(-0.49)	(-2.93)	(3.48)	(3.28)
pgdp		1.135***		-0.431**		1.016***		0.261
		(14.25)		(-2.51)		(4.49)		(1.22)
gov		0.003		0.016		0.420***		0.009
		(0.06)		(0.14)		(2.84)		(0.07)
env		-0.012		0.332***		0.379***		-0.215**
		(-0.31)		(3.85)		(3.34)		(-2.00)
scale		0.079		-0.678**		-1.385***		-0.036
		(0.61)		(-2.45)		(-3.80)		(-0.10)

Variable	TC		TEC		SC		AEC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ump		0.022		0.310***		0.113		0.045
		(0.80)		(5.17)		(1.43)		(0.60)
isT		-0.099***		-0.090		0.312***		-0.410***
		(-2.89)		(-1.22)		(3.22)		(-4.48)
eel		0.028		-0.139**		0.153**		-0.094
		(1.05)		(-2.44)		(2.05)		(-1.33)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.712	0.886	0.506	0.591	0.073	0.209	0.289	0.335
Sample size	399	399	399	399	399	399	399	399

According to the estimation results in Table 6, columns (1) and (2) show that the estimated coefficient of GF on TC is 0.108, which is statistically significant and positive at the 1% confidence level. This indicates that the development of green finance has promoted technological progress in Guangdong Province, consistent with hypothesis H2a. This suggests that green finance development has successfully guided capital and resources to the construction of green-related fields in the region, improved the availability of funds for enterprises, and driven corporate investment in green technology through related funding and policies, thereby stimulating regional technological innovation and ultimately enhancing overall technological progress in the region.

From columns (3) and (4), the estimated coefficient of GF on TEC is -0.142, which is statistically significant and negative at the 1% confidence level. This demonstrates that green finance development has a negative impact on technical efficiency in Guangdong Province, aligning with hypothesis H3b. A possible explanation is that green finance development has redirected capital and resources to energy-saving and environmentally friendly enterprises, leaving non-environmentally friendly enterprises—particularly small and medium-sized ones—struggling to obtain sufficient funding. Consequently, these enterprises face greater difficulties in achieving and acquiring innovative technologies, while simultaneously experiencing increased management costs, R&D expenses, and risks. Moreover, under green development pressures, non-environmentally friendly enterprises are more susceptible to external financing constraints, leading to declining technical efficiency and ultimately affecting regional technical efficiency improvements.

Columns (5) and (6) reveal that the estimated coefficient of GF on SC is negative and becomes statistically significant at the 1% confidence level after controlling for other variables. This indicates that green finance development has a significant negative impact on scale efficiency in Guangdong Province, supporting hypothesis H4b. The rapid advancement of green finance development may have channeled resources to already saturated enterprises, causing them to deviate from optimal scales. Simultaneously, non-environmentally friendly enterprises receive less support and experience lower production efficiency. Given Guangdong's historical context of extensive economic expansion and massive influx of talent, enterprises, and resources, green finance development may have exacerbated these issues.

Columns (7) and (8) show that the estimated coefficient of GF on AEC is statistically significant and positive at the 1% confidence level, indicating that green finance development has positively influenced resource allocation efficiency in Guangdong Province, consistent with hypothesis H5a. This demonstrates that green finance development has redirected capital and resources from low-efficiency enterprises to high-efficiency ones, and from polluting enterprises to non-polluting ones. These resource reallocations have generated positive effects on total factor productivity in the region.

4.3 Robustness Test

This study conducted robustness checks by replacing the dimensionless normalization method in the GF indicator measurement from the normalization approach to standardization, recalculating GF as GF2, and re-estimating the regression model. Table 7 results demonstrate that GF after the measurement method replacement still exerts a statistically significant negative impact on TFPC, further verifying the reliability of hypothesis H1b.

Table 7. Robustness test results

Variable	TFPC							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GF2	0.036	-0.060	-0.206***	-0.206***	-0.201***	-0.192***	-0.192***	-0.185***
	(0.64)	(-1.05)	(-2.88)	(-2.90)	(-2.91)	(-2.78)	(-2.79)	(-2.69)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.135	0.197	0.221	0.230	0.272	0.279	0.288	0.295
Sample size	399	399	399	399	399	399	399	399

4.4 Endogeneity Test

Despite rigorous efforts to control confounding factors, persistent risks of omitted variable bias and potential reverse causality result in endogeneity concerns. This study employs an instrumental variable (IV) approach, drawing on methodological frameworks from Guo, Zhang & Liu (2022), by utilizing Internet penetration rate (iba) operationalized through broadband subscriber density. Given the presence of serial correlation in total factor productivity measurements, we implement a system GMM dynamic panel model with dual fixed effects to address this methodological challenge and verify our baseline findings. The regression outputs presented in Table 8 reveal significantly negative coefficients for GF, mirroring both magnitude and significance from our baseline specifications, demonstrating that green finance development's adverse impact on regional productivity remains statistically robust.

Table 8. SYS-GMM estimation results

Variable	Phase 1 results GF	Phase 2 results TFPC
Iba	0.262***	
	(2.98)	
GF		-0.933**
		(-2.05)
Control variables	Yes	Yes
Time fixed effects	Yes	Yes
Individual fixed effects	Yes	Yes
R^2	0.395	0.379
Sample size	399	399

5. Conclusions and Policy Implications

5.1 Conclusions

To advance the coordinated development of green finance and the economy in Guangdong Province, this paper employs panel data from 21 prefecture-level cities (2000–2019) to empirically analyze the mechanisms through which green finance affects regional total factor productivity (TFP), with a focus on technological progress, technical efficiency, scale efficiency, and allocative efficiency. The findings are as follows: (1) Green finance development in Guangdong has a significant negative impact on regional TFP growth. This conclusion remains robust after addressing endogeneity concerns and recalculating the green finance indices. (2) Mechanism analysis reveals that while green finance promotes technological progress and improves resource allocation efficiency, it simultaneously hinders improvements in technical and scale efficiency. The net effect on regional TFP is negative. This indicates that the rapid advancement of green finance exerts a complex influence on TFP. (3) In terms of the internal impact structure, the suppression of TFP growth by green finance is primarily attributable to its negative effects on regional technical and scale efficiency.

5.2 Policy Recommendations

First, policies should be tailored to different types of enterprises. Priority must be given to transforming traditional, non-energy-efficient, and environmentally unfriendly enterprises by imposing stricter penalties on underperformers. For companies with opaque or problematic environmental disclosures, appropriate sanctions should be applied, and oversight of their green financial assets must be strengthened to enhance the quality and sustainability of regional green finance. Increasing the strength and coverage of green policy support, particularly by lowering green credit interest rates, will reduce financing, transformation, and R&D costs for traditional firms. This will boost their willingness to innovate and transform, stimulate regional technological innovation, improve the liquidity and allocation efficiency of green financial resources across traditional and green industries, and ultimately promote industrial upgrading and regional scale efficiency.

Second, green finance policies should be regionalized. Given the varying levels of economic development across Guangdong, greater supervision over fund and resource allocation should be exercised in less developed areas. This will ensure that even amid unavoidable rough economic expansion, enterprises use resources more prudently and efficiently. At the same time, differentiated green economic development programs should be designed for different regions to ensure effective implementation and avoid a "one-size-fits-all" approach. By deeply understanding local conditions and needs, we can formulate green finance policies that are better aligned with reality. Moreover, inter-regional cooperation in green finance should be strengthened, encouraging developed areas to collaborate with less developed ones. Establishing a cross-regional cooperation mechanism involving financial institutions, enterprises, and governments can facilitate technology transfer and experience sharing, thereby raising regional technical efficiency and mitigating the diminishing efficiency returns that often accompany economic expansion.

Third, policies should target the specific pathways through which green finance affects total factor productivity. Programs supporting green technological innovation should be established to help enterprises secure more funding for adopting high-efficiency technologies, upgrading low efficiency processes, and developing green technologies and products. Financial institutions should also be encouraged to innovate in green financial products. These measures will elevate the region’s overall technological level, amplify the role of technological progress in boosting productivity, and improve technological efficiency.

Additionally, green support and regulation should be enhanced, information transparency improved, and enterprises encouraged to allocate resources more rationally. This will increase capital and resource allocation efficiency across industries and magnify its positive impact on total factor productivity. By refining the green finance evaluation system to assess and rank the practices of institutions and enterprises, we can incentivize continuous improvement in the quality and scale of green business and bolster policy support for green transformation and production scaling. This will help low efficiency enterprises optimize their structure, reduce costs, and improve technological and allocation efficiencies. Simultaneously, the green finance risk management system should be upgraded to accurately assess project level environmental risks, enabling financial institutions to correctly identify and price such risks. This will mitigate the potential negative impact of risks on technical efficiency and help maintain regional technical efficiency at a relatively stable level. Finally, public and corporate

environmental awareness should be raised through media campaigns, education, and training. The government should also refine green planning and avoid excessive economic expansion to achieve sustainable gains in scale efficiency.

Funding

This research was generously supported by the National Natural Science Foundation of China (Grant No. 42461027) and the Guangdong Office of Philosophy and Social Science (Grant No. Grant No. GD24XLJ06 & GD24YYJ09).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Alawi, S. M., Abbassi, W., Saqib, R., & Sharif, M. (2022). Impact of financial innovation and institutional quality on financial development in emerging markets. Journal of Risk and Financial Management, 15(3), 1-11.

[2] Al-Yahya, S. N. (2022). The COVID-19 pandemic uncertainty, investor sentiment, and global equity markets: Evidence from the time-frequency co-movements. The North American Journal of Economics and Finance.

[3] Aslam, B., Zhang, G., Amjad, M. A., Guo, S., & Ji, M. (2023). Does the impact of financial development reinforce sustainability ecological footprint? Fresh evidence from middle and high-income economies. Journal of Cleaner Production, 429.

[4] Battese, C. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency, and panel data: With application to paddy farmers in India. Journal of Productivity Analysis, 3(1-2), 153-169.

[5] Chavriya, S., Sharma, G. D., & Mahendru, M. (2024). Financial inclusion as a tool for sustainable macroeconomic growth: An integrative analysis. Annals of Public and Cooperative Economics, 95(2).

[6] De Haas, R., & Popov, A. (2023). Finance and green growth. The Economic Journal, 133(650), 637-668.

[7] Falcone, P. M. (2020). Environmental regulation and green investments: The role of green finance. International Journal of Green Economics, 14(2), 159-173.

[8] Goldstein, I., Koijen, R. S. J., & Mueller, H. M. (2021). COVID-19 and its impact on financial markets and the real economy. The Review of Financial Studies, 34(11), 5135-5148.

[9] Gu, B., Chen, F., & Zhang, K. (2021). The policy effect of green finance in promoting industrial transformation and upgrading efficiency in China: Analysis from the perspective of government regulation and public environmental demands. Environmental Science and Pollution Research, 28, 47474–47491.

[10] Guo, J., Zhang, K., & Liu, K. (2022). Exploring the mechanism of the impact of green finance and digital economy on China's green total factor productivity. International Journal of Environmental Research and Public Health, 19(23), 16303.

[11] Guo, X., Yang, J., Shen, Y., & Zhang, X. (2024). Impact of green finance and environmental regulation on carbon emissions: Evidence from China. Frontiers in Environmental Science, 12, 1307313.

[12] He, L., Zhong, T., & Gan, S. (2022). Green finance and corporate environmental responsibility: Evidence from heavily polluting listed enterprises in China. Environmental Science and Pollution Research, 29, 74081–74096.

[13] Hou, H., Chen, M., & Zhang, M. (2023). Study on high energy-consuming industrial agglomeration, green finance, and carbon emission. Environmental Science and Pollution Research, 30, 29300–29320.

[14] Hu, K., Kuang, W., & Qin, Q. (2024). Advancing resource efficiency through renewable energy, green innovations, and education: Lessons from G20 economies. Journal of the Knowledge Economy.

[15] Kirci Altinkeski, B., Cevik, E. I., & Dibooglu, S. (2022). Investor sentiments and stock markets during the COVID-19 pandemic. Financial Innovation, 8, Article number: 69.

[16] Kumbhakar, S. C., & Lovell, C. A. K. (2000). Stochastic frontier analysis. Cambridge University Press.

[17] Lei, H., Gao, R., Ning, C., & Sun, G. (2024). Impact of green finance on regional green innovation performance. Sustainability, 16(23), 10519.

[18] Leibenstein, H. (1966). "A-efficiency" vs "X-efficiency." International Economic Review, 7(2), 56.

[19] Li, J., & Zhang, L. (2023). Green innovation and consumer demand for sustainable products: Evidence from China.

Journal of Cleaner Production, 354.

[20] Lin, G., Lam, J. F. I., Shi, Y., Chen, H., & Chen, H. (2023). Revisiting the linkage between green finance and China's sustainable development: Evidence from the pilot zones for green finance reform innovations. Frontiers in Ecology and Evolution, 11, 1264434.

[21] Liu, W., & Zhu, P. (2024). The impact of green finance on the intensity and efficiency of carbon emissions: The moderating effect of the digital economy. Frontiers in Environmental Science, 12, 1362932.

[22] Luo, Q., & Wang, J. (2024). Digital finance and green technology innovation: A dual path test based on market and government. Finance Research Letters, 70, 106283.

[23] Maltais, A., & Nykvist, B. (2021). Understanding the role of green bonds in advancing sustainability. Journal of Sustainable Finance & Investment, 11(3), 233-252.

[24] Matteo, P., Graziana, G., Matilda, S., et al. (2023). Green finance: Past, present and future studies. Journal of Financial Management, Markets and Institutions, 11(02).

[25] Oliveira, H., Moutinho, V. F., & Afonso, O. (2024). Exploring the relationship between technological progress, human capital, political uncertainty, energy consumption, and economic growth: Evidence from a panel data analysis. Journal of Knowledge Economy.

[26] Palmaccio, M., Galeone, G., Shini, M., & Campobasso, F. (2023). Green finance: Past, present and future studies. Journal of Financial Management, Markets and Institutions (JFMMI), 11(02), 1-19.

[27] Rasoulinezhad, E., & Taghizadeh-Hesary, F. (2022). Role of green finance in improving energy efficiency and renewable energy development. Energy Efficiency, 15(2), 14.

[28] Rupeika-Apoga, R., & Petrovska, K. (2022). Barriers to sustainable digital transformation in micro-, small-, and medium-sized enterprises. Sustainability, 14(20), 13558.

[29] Sajid, M., Ansari, M. A. A., Tanveer, A., & et al. (2023). Evaluating the influence of green growth, institutional quality, and financial inclusion on financial stability: Evidence by sustainable finance theory. Environmental Science and Pollution Research, 30, 115965–115983.

[30] Wang, H., & Shao, R. (2024). Does the green finance policy affect the efficiency of corporate investment? Evidence from China's green finance reform and innovation pilot zones. PLoS ONE, 19(11), e0313861.

[31] Wang, Q., & Fan, Z. (2023). Green finance and investment behavior of renewable energy enterprises: A case study of China. International Review of Financial Analysis, 87.

[32] Wei, N. (2024). Green finance, market integration, and regional economic resilience. Finance Research Letters(PA), 105777-105777.

[33] Wu, G., Xu, Q., Niu, X., & Tao, L. (2022). How does government policy improve green technology innovation: An empirical study in China. Frontiers in Environmental Science, 9, 799794.

[34] Yang, Z. (2023). The impact of green finance on high-quality economic development in China: Vertical fiscal imbalance as the moderating effect. Sustainability, 15(12), 9350.

[35] Zeng, Q., Tong, Y., & Yang, Y. (2023). Can green finance promote green technology innovation in enterprises: Empirical evidence from China. Environmental Science and Pollution Research, 30(3), 1-14.

[36] Zhang, C., Wang, C., Jin, H., Xu, M., Balezentis, T., & Agnusdei, L. (2024). The impact of green finance on energy transition and carbon emission targets: Exploring the double threshold and spatial spillover effects among the regions. Structural Change and Economic Dynamics.

[37] Zhang, D., Wang, C., Miao, S., & Deng, L. (2024). The impact of firm's ESG performance on the skill premium: Evidence from China's green finance reform pilot zone. International Review of Financial Analysis, 89, 102549.

[38] Zhang, J., Wu, G., & Zhang, J. (2004). Interprovincial physical capital stock estimation in China: 1952–2000. Economic Research, 10, 35–44.

[39] Zhang, W., Ke, J., Ding, Y., & et al. (2024). Greening through finance: Green finance policies and firms' green investment. Energy Economics, 131, 107401.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Impact of Enterprise Digitalization on the Governance of Multi-Partner R&D Alliances

Jinlai Chen, Jie Liang*, Shuhan Meng

School of Management, Xi'an Polytechnic University, Xi'an, Shaanxi, 710048, China

Corresponding author: Jie Liang, liangjie@xpu.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Addressing the limitation of existing research that rarely focuses on the governance of digital transformation and R&D alliances, this study explores the impact of digital transformation on the number of new partners in R&D alliances. Based on an empirical analysis of 902 samples from 133 listed pharmaceutical manufacturing companies in Shanghai and Shenzhen A-share markets from 2011 to 2022, the results show that the degree of digital transformation promotes the number of new partners in R&D alliances, while the speed of digital transformation hinders it. These findings develop the theoretical framework for existing R&D alliance governance factors and deepen the academic understanding of the essence of digital transformation.

Keywords: Relationship Governance; Digital Transformation; Number of Relationships

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1408

1. Introduction

The digital economy has become a core engine of global economic development and a key area and important engine for accelerating the development of new quality productivity. With the accelerated iteration and deep integration of digital technologies such as big data, artificial intelligence, and cloud computing, the business environment is undergoing profound changes characterized by digitalization. The comprehensive advancement of the Digital China strategy requires enterprises to widely apply new technologies such as informatization, digitalization, and artificial intelligence to accelerate their digital transformation. Against this macro-background, digitalization is no longer a strategic option for enterprises, but a crucial question concerning their survival and development.

Digitalization has not only reshaped internal production processes and management models within enterprises, but has also profoundly broken down traditional organizational boundaries, enabling companies to integrate resources and collaborate in a broader space. Theoretically, enhanced digital capabilities should help companies reduce collaboration search costs, improve information processing capabilities, and expand partner networks, thus providing strong support for the construction and expansion of R&D alliances. However, the reality presents a complex picture. In recent years, Chinese enterprises have frequently encountered governance challenges in collaborative innovation practices, such as inconsistent goals and difficulties in coordinating cooperative relationships, resulting in a persistently high failure rate for multi-partner R&D alliances. At the same time, enterprises still lack experience in operating R&D alliances; patents jointly applied for by multiple organizations account for only about 20%, reflecting that the penetration and maturity of multilateral collaborative innovation in Chinese

enterprise practices still need to be improved.

R&D alliances, as a crucial pathway for corporate innovation, are fundamentally about resource sharing, information exchange, and collaborative innovation among different companies based on shared technological goals. With the advancement of globalization, companies face increasing pressure to innovate, and the R&D capabilities of a single company are no longer sufficient to meet innovation demands. Therefore, more and more companies are choosing to form R&D alliances through collaboration with external companies, research institutions, and other stakeholders. The quality of relationships within an R&D alliance directly determines its innovative effectiveness. Research shows that good cooperative relationships effectively promote knowledge sharing, information flow, and technological innovation, while poor cooperative relationships can lead to resource waste, communication barriers, and even conflicts. Therefore, how to effectively manage and improve the quality of relationships within R&D alliances has become an important issue of concern for both academia and industry.

The theoretical contributions of this study include the following aspects. First, it deconstructs enterprise digitalization into two dimensions: the degree of digitalization and the speed of digitalization. This breaks through the limitations of previous studies that mostly focused on a single dimension, revealing the differentiated impact of different dimensions of digitalization on the expansion of R&D alliances and deepening the theoretical understanding of the connotation of digitalization. Second, it expands digitalization research from an internal enterprise perspective to an inter-organizational relationship perspective, exploring how digitalization affects the construction of external collaborative networks. Finally, it provides a new analytical framework for understanding the relationship between digitalization and innovation, revealing the “double-edged sword” effect of digitalization on the expansion of collaborative networks—the coexistence of the empowering effect of the degree of digitalization and the inhibitory effect of the speed of digitalization, which helps to reconcile the divergent conclusions in existing research on the effects of digitalization.

2. Literature Review

2.1 Digital Transformation

Existing research primarily discusses the connotation of digital transformation from two perspectives: enterprise technology and organizational change. From a technology perspective, scholars generally view digital transformation as a process of applying and integrating a series of advanced digital technologies, with technology as the core driving engine. Studies by Bharadwaj et al. argue that enterprises utilize digital technology resources such as big data, artificial intelligence, and cloud computing to build new digital capabilities and develop digital products and services. From an organizational change perspective, the focus shifts from the technological tools themselves to how technology triggers deep structural, process, and cultural changes within enterprises [1]. Kane et al. pointed out that digital transformation is far from a simple technological upgrade, but rather a process in which enterprises carry out comprehensive innovation in organizational structure, business processes, talent skills and even organizational culture in order to cope with the disruptive changes in the digital environment [2]. At the same time, some scholars tend to combine the perspectives of technology and organization to form a systematic understanding. Vial et al. believe that digital transformation is a process that triggers major changes through digital technology, aiming to cope with disruptive changes in the environment by improving existing or creating new value propositions, operating models, customer experiences and relationships [3].

Digital transformation can be further divided into static digital transformation degree and dynamic digital transformation speed. Existing research on enterprise digital transformation largely focuses on the static digital transformation degree. These studies suggest that digital transformation can enhance an organization’s search and learning capabilities by expanding its search scope, increasing the amount of information and knowledge acquired, accelerating the processing speed of information and knowledge, and increasing the possible combinations of information and knowledge [4]; Furthermore, digital transformation can also leverage digital technologies to integrate partners into digital platforms, reducing the costs and risks of inter-organizational collaboration and strengthening cooperation [5].

Digital transformation is a dynamic and evolving process, and the speed of transformation is also a key characteristic. Enterprises need to find the optimal transformation speed based on their own resources and capabilities, balancing the

pressure of technological catch-up with the costs of organizational adaptation [6]. Existing research indicates that while a faster pace of digital transformation accelerates a company's digital transformation process, the significant investment of capital, human resources, and effort can crowd out resources from other business operations[7]. At the same time, rapid transformation in the short term poses significant challenges to the coordination between digital resources and capabilities and the organization's original resources and capabilities, as well as the coordination between new organizational rules under digitalization and existing organizational rules. This may lead to an imbalance in the dynamic capabilities of enterprise organizational management[8].

Regarding the consequences of digital transformation, existing literature mainly studies its impact on corporate performance and innovation capabilities. Wang et al.'s research found that digital transformation improves corporate performance by reducing costs, increasing R&D intensity, and enhancing human resources, and has a significant positive impact on the overall manufacturing industry [9]. Xu et al.'s research indicates that digital transformation significantly improves enterprises' innovation performance, and the dual mediating effect of big data capabilities and organizational agility is an important part of its influencing mechanism [10].

2.2 Number of R&D Alliance Relationships

A research and development alliance (R&D alliance) refers to a formal, but not merged, cooperative organization formed by two or more enterprises and related institutions based on market opportunities and the need for independent innovation, with the aim of jointly developing new technologies and products. This alliance is characterized by complementary strengths, shared benefits, and shared risks. Depending on the name, R&D alliances include R&D alliances industry-academia-research alliances, technology alliances, knowledge alliances, patent alliances, technology standards alliances, joint R&D efforts, and virtual R&D alliances, etc.[11]. Since the 1990s, some research has focused primarily on the issue of firms forming single R&D alliances from a bilateral perspective. During this phase, scholars mainly focused on the impact of individual partner characteristics on innovation performance, including the partner's type, resources, relationships, and geographical location [12]. Since the 1990s, research on the governance effects of corporate innovation networks has attracted much attention from scholars. Among them, the theoretical perspective based on network relationships is an important entry point for the academic community to explain corporate innovation networks and outputs [13]. This research approach often approaches innovation performance from the perspective of alliance governance, exploring the mechanisms by which relationships among participants influence innovation performance [14]. These studies on network relationships primarily focus on the degree of development of cooperative relationships among partners, specifically including three dimensions: firstly, the number of cooperative relationships [15]. Secondly, the strength of the relationship, including intimacy, familiarity, trust, relationship embedding, and frequency of interaction [16]. The third is the fault of the relationship [17]. Research on the relationship between network relationship governance and innovation performance has yielded inconsistent conclusions, suggesting that partnerships are a "double-edged sword," potentially promoting or hindering innovation under different conditions.

3. Theoretical Hypothesis

3.1 The Degree of Digital Transformation and the Number of New R&D Alliance Partners

This study elucidates the impact of digital transformation on the number of new partners in corporate R&D alliances from three aspects.

First, digitalization enhances a company's information processing capabilities, reducing partner search costs. Highly digitalized companies can leverage big data analytics, artificial intelligence algorithms, and other technological tools to identify potential partners in a broader context, assessing their technological capabilities, resource endowments, and reputation for cooperation, thereby significantly reducing information asymmetry and search costs. Second, digitalization strengthens a company's resource integration capabilities, increasing its attractiveness to new partners. Companies with higher levels of digitalization often possess stronger resource integration capabilities and technological compatibility, providing potential partners with more convenient interfaces and more efficient collaboration platforms. The construction of digital platforms allows companies to incorporate partners into a unified technological framework, reducing coordination costs and technological barriers in cross-organizational collaboration. This platform capability makes highly digitalized companies more attractive partners in R&D alliance networks,

attracting more partners to actively seek cooperation. Finally, digitalization promotes corporate knowledge transparency, strengthening the foundation of trust in cooperation. The application of digital technologies allows information such as a company's technological capabilities, R&D direction, and cooperation history to be presented more transparently, reducing the uncertainty of potential partners' perception of a company's capabilities. This transparency effect helps to quickly establish initial trust in the early stages of cooperation, laying the foundation for deeper collaboration later. Research on alliance and combination networks indicates that resource flow plays a positive mediating role between network connectivity and innovation capabilities, and the degree of digitalization is a key factor promoting resource flow. Based on the above analysis, this study proposes the following hypotheses.

H1: The degree of digital transformation contributes to the number of new partners in the R&D alliance.

3.2 Speed of Digital Transformation and Number of New Partners in R&D Alliances

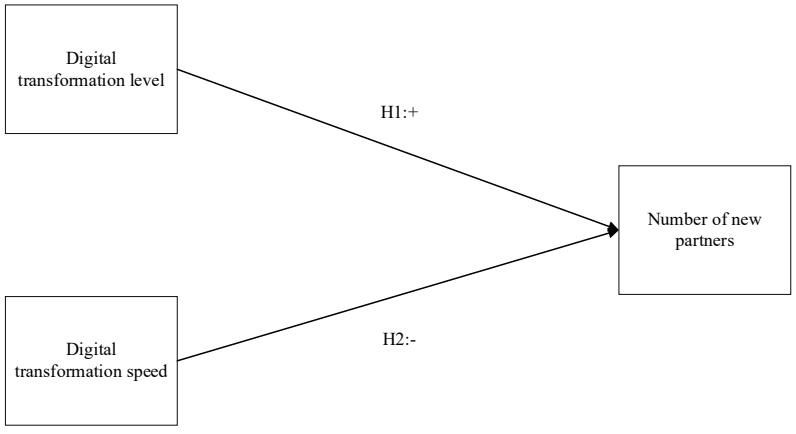
This study elucidates the impact of the speed of digital transformation on the number of new partners in corporate R&D alliances from three aspects.

First, rapid transformation leads to a resource crowding-out effect, dispersing resources needed for expanding cooperation. A faster pace of digital transformation means that companies need to invest significant funds, human resources, and management effort in technology upgrades, process reengineering, and talent development in the short term. These resource investments are highly exclusive, potentially crowding out resources that could be used to expand new partnerships and maintain existing ones. When companies focus primarily on internal digital transformation, their attention to and management of external cooperation networks inevitably decreases, thus inhibiting the establishment of new partnerships. Second, rapid transformation exacerbates internal organizational turmoil and reduces the stability of external cooperation. Excessive speed of digital transformation can lead to internal process turmoil, restructuring of authority and responsibility systems, and changes in personnel skills. These organizational-level instabilities will be transmitted to the external cooperation level. Potential partners may have doubts about the performance capabilities, technological compatibility, and sustainability of cooperation when establishing partnerships with companies undergoing rapid transformation, thus reducing their willingness to cooperate. Research on alliance management capabilities shows that effective alliance management requires a stable organizational foundation and continuous resource investment, which rapid transformation erodes. Finally, rapid transformation may lead to a lag in organizational learning capabilities, making it difficult to effectively absorb external knowledge. Dynamic capability theory suggests that the speed of organizational change should match its learning capacity. When the speed of digital transformation exceeds the organization's learning capacity, enterprises will fall into a "digestion" dilemma, struggling to effectively absorb and integrate external knowledge resources. In this situation, even if enterprises encounter new partners, they will find it difficult to obtain substantial innovative benefits from the collaboration, leading to superficial and short-term partnerships, which is detrimental to establishing sustainable and stable R&D alliances. Based on the above analysis, this study proposes the following hypothesis:

H2: The speed of digital transformation hinders the number of new partners in R&D alliances.

Figure 1 is the research model diagram of this paper.

Figure 1 Research model



4. Research Design

4.1 Samples and Data Sources

This study selected listed companies in China's pharmaceutical manufacturing industry from 2011 to 2022 as research subjects. The pharmaceutical manufacturing industry is characterized by its technology innovation orientation and R&D alliance orientation , This research is suitable for studying issues related to R&D alliances, and many existing studies on corporate R&D alliances are based on the pharmaceutical manufacturing industry [19]. This study determines the formation of R&D alliances and the selection of partners based on enterprise patent data [20]. The required patent data was obtained from the China National Intellectual Property Administration. The study also retrieved historical financial information and corporate governance data from company websites and the Guotai An database. Regarding data processing, this study first eliminated ST/PT companies and samples with excessive missing data based on the list of listed pharmaceutical manufacturing companies published by the China Securities Regulatory Commission to ensure data completeness and reliability. Secondly, it identified other organizations that had jointly applied for patents with the aforementioned companies within a certain period, using patent data disclosed by the China National Intellectual Property Administration, to establish R&D alliances [21]. The required patent data was obtained from the China National Intellectual Property Administration. The study also retrieved historical financial information and corporate governance data from company websites and the Guotai An database. Regarding data processing, this study first eliminated ST/PT companies and samples with excessive missing data based on the list of listed pharmaceutical manufacturing companies published by the China Securities Regulatory Commission to ensure data completeness and reliability. Secondly, it identified other organizations that had jointly applied for patents with the aforementioned companies within a certain period, using patent data disclosed by the China National Intellectual Property Administration, to establish R&D alliances. [18]. Then, all continuous variables were shortened at the 1st and 99th percentiles to address outlier issues. Finally, considering the causal relationship between the independent and dependent variables, the variables were lagged by one period to address reverse causality. After these processing steps, this study obtained a sample of 902 R&D alliances from 133 listed companies in China's pharmaceutical manufacturing industry from 2011 to 2022.

4.2 Variable measurement

4.2.1 Explained Variable

Number of new R&D alliance partners (NEW). The selection of new R&D alliance partners refers to a company's choice of partners who possess heterogeneous knowledge resources, cutting-edge technological capabilities, and a high willingness to take risks [22]. Existing research typically uses new partners with whom one has never collaborated before as a proxy for new partner selection [23] Therefore, this study also adopts this approach, using the number of other organizations with which firms have no prior patent collaborations as a measure of the degree of exploratory partner selection. Specifically, this study counts the number of new partners with which firms engage in patent collaboration for the first time within a five-year window of a specific year. Considering the potential for significant bias in the final data, to ensure the robustness of the results, this study adds 1 to the number of new partners and then takes the logarithm [24].

4.2.2 Explanatory Variables

Digital Transformation Degree (DT): Referring to the research of Li Yingmei et al [25], this paper uses the logarithm of the enterprise's digital intangible assets to measure the degree of enterprise digital transformation. Specifically, the first step is to refer to existing research to determine the keyword "digitalization" and identify the digital part of the enterprise's intangible assets through this digital keyword. That is, according to the name of the detailed account of "intangible assets" in the financial statements of listed companies, the digital keyword is screened, and the investment is the investment related to digitalization. Digital Transformation Speed (DTV): Based on the approach of Du Shanzhong and Ma Lianfu [26] , this paper also uses the above-mentioned degree of enterprise digital transformation as the basis and adopts relative speed to measure the speed of enterprise digital transformation, so as to more directly reflect its changing trend.

4.2.3 Control Variables

Referring to the existing research, this study selects enterprise age, enterprise size, R&D expenses, partner type, equity

concentration, market-oriented index, operating income growth rate, and debt ratio as control variables.

4.3 Model Building

To investigate the impact of digital transformation on the number of new partners in R&D alliances, this study focuses on listed pharmaceutical manufacturing companies in Shanghai and Shenzhen, China, and employs a multidimensional fixed effects model for estimation, specifically a dual fixed effects model that controls for time trends and individual factors, for data analysis. Robust standard errors were assumed in the regression tests.

5. Results

5.1 Descriptive Statistics

Table 1 shows the mean, standard deviation, and correlation coefficient analysis results of the variables in this study. The absolute values of the correlation coefficients among all variables are less than 0.5, indicating that high correlation between variables will not lead to estimation bias. Furthermore, the VIF (Variance Inflation Factor) values for all variables are less than 10, indicating that there is no serious multicollinearity problem among the variables.

5.2 Benchmark Regression

Table 2 presents the regression results of the degree and speed of digital transformation on the number of new partners in the R&D alliance.

Hypothesis 1 proposes that the degree of digital transformation promotes the number of new partners in the R&D alliance. The first column of Table 2 shows that the degree of digital transformation has a significant positive effect on the number of new partners in the R&D alliance; that is, as the degree of digital transformation of an enterprise increase, it can attract more new partners to join the R&D alliance. Thus, Hypothesis 1 is true.

Hypothesis 1 proposes that the speed of digital transformation hinders the number of new partners in the R&D alliance. The second column of Table 2 shows that the speed of digital transformation has a significant negative effect on the number of new partners in the R&D alliance, meaning that excessively rapid digital transformation may have a limiting effect on attracting new partners. Thus, Hypothesis 2 is true.

Table 1 Descriptive statistics and correlation coefficient analysis

	NEW	DT	DTV	Fage	Size	Rd	Partner type	Market	Growth	Lev	Top10
NEW	1.000										
DT	0.209***	1.000									
DTV	-0.038	0.128***	1.000								
Fage	0.041	0.243***	-0.024	1.000							
Size	0.207***	0.554***	0.057*	0.226***	1.000						
Rd	0.141***	0.068**	0.063*	0.167***	0.215***	1.000					
Partner type	0.658***	0.122***	-0.002	0.110***	0.212***	0.028	1.000				
Market	0.140***	0.198***	0.041	0.215***	0.210***	0.155***	0.179***	1.000			
Growth	-0.010	-0.029	0.063*	-0.095***	-0.036	-0.026	0.001	0.005	1.000		
Lev	-0.028	0.270***	0.022	0.124***	0.319***	-0.044	0.018	0.037	-0.006	1.000	
Top10	0.015	0.016	0.077**	-0.150***	0.108***	0.059*	-0.094***	-0.130***	0.109***	-0.172***	1.000
Mean	1.971	14.606	0.016	2.917	7.880	17.306	1.261	9.112	0.164	0.322	57.688
SD	1.863	1.504	0.526	0.304	0.963	18.891	0.848	1.971	0.278	0.164	13.359

Note: Correlation coefficients with absolute values higher than 0.057 are significant at a two-tailed test of 0.1, and correlation coefficients with absolute values higher than 0.094 are significant at a two-tailed test of 0.01. N=902.

Table 2 Benchmark effects

	(1) NEW	(2) NEW
DT	0.078(0.045)**	
DTV		-1.740(0.857)**
Control variables	Include	Include
Year mute variable	Include	Include
cons	7.617(3.861)**	8.096(3.817)**
F	33.956	34.181
adj. R^2	0.623	0.624

Note: Standard errors in parentheses. Hypothesis test coefficients are one-tailed, control variable coefficients are two-tailed; * p<0.1, ** p<0.05, *** p<0.01. N=902.

6. Conclusion and Discussion

Given the limited scope of existing research on digital transformation and the governance of R&D alliance relationships, this study analyzes the impact of digital transformation on the number of new partners in R&D alliances. The results show that (1) as the degree of digital transformation increases, more new partners can be attracted to join the R&D alliance. (2) Overly rapid digital transformation may have a limiting effect on attracting new partners.

The main contributions of this paper are as follows: First, it expands the research perspective on digital transformation, extending it from the internal performance level of enterprises to the external cooperation network level, revealing the promoting effect of the degree of digital transformation on the number of new partners in R&D alliances. Second, it distinguishes between the two dimensions of transformation degree and transformation speed, finding that they have opposite effects, enriching the theoretical connotation of the dynamic process of digital transformation. Finally, this study provides a reference for enterprises to formulate the pace and path of digital transformation, emphasizing that while increasing the depth of transformation, the speed of advancement should be reaonably controlled to achieve a steady expansion of the cooperation network.

This study also has certain limitations. First, the sample data mainly comes from certain specific industries and regions and may not fully represent the situation of enterprises in all industries and regions. Secondly, the study only focused on the single dimension of the degree and speed of digital transformation, without exploring other factors that may influence the number of new partners in R&D alliances, such as corporate culture and management models. Future research could consider collecting data from a broader scope to further enrich the theoretical model and explore the impact of other variables on partner selection.

Funding

Shaanxi Provincial Innovation Capability Support Program Project: Research on the Mechanism and Countermeasures of Digitalization Affecting the Governance of Multi-Partner R&D Alliance Relationships and Green Innovation Performance (2025KG-YBXM-143)

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: Toward a next generation of insights. MIS Quarterly, 37(2), 471–482.

[2] Kane, G. C., Palmer, D., Phillips, A. N., & Kiron, D. (2015). Is your business ready for a digital future? MIT Sloan Management Review, 56(4), 37.

[3] Vial, G. (2021). Understanding digital transformation: A review and a research agenda. In Managing digital transformation (pp. 13–66). Springer.

[4] Qian, H., Xiao, J., Li, S., & Wang, Y. (2024). Digital transformation and enterprise innovation: Evidence from the Chinese biomedical industry. Industrial Economics Research, (05), 114–127.

[5] Fang, M., Lin, J., & Nie, H. (2022). Does digital transformation promote common prosperity within firms? Evidence from Chinese A-share listed firms. Journal of Quantitative & Technological Economics, 39(11), 50–70.

[6] Sun, B., Zhang, Y., Zhu, K., & Li, X. (2024). Is faster really better? The impact of digital transformation speed on firm financial distress: Based on the cost-benefit perspective. Journal of Business Research, 179, 114703.

[7] Kesidou, E., Ri, A., & Roper, S. (2025). Balancing act or two roads to travel: Evaluating the trade-offs between digitalization and net zero innovation in SMEs. Small Business Economics, 1–30.

[8] Li, Y., & Meng, X. (2024). How does the optimization of the business environment affect the digitalization paradox of enterprises? ——From the perspective of innovation persistence. Finance and Accounting Monthly, 45(06), 20–27.

[9] Wang, Y., Wang, T., & Wang, Q. (2024). The impact of digital transformation on enterprise performance: An empirical analysis based on China's manufacturing export enterprises. PLOS ONE, 19(3), e0299723.

[10] Xu, M., Zhang, Y., Sun, H., & Liu, J. (2024). How digital transformation enhances corporate innovation performance: The mediating roles of big data capabilities and organizational agility. Heliyon, 10(14).

[11] Huang, S., Cong, J., & Chao, H. (2016). Organization and governance of research and development alliances: A literature review and future prospects. Foreign Economics & Management, 38(06), 63–81+99.

[12] Capaldo, A., & Petruzzelli, A. M. (2014). Partner geographic and organizational proximity and the innovative performance of knowledge-creating alliances. European Management Review, 11(1), 63–84.

[13] Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. American Sociological Review, 674–698.

[14] Grimpe, C., & Kaiser, U. (2010). Balancing internal and external knowledge acquisition: The gains and pains from R&D outsourcing. Journal of Management Studies, 47(8), 1483–1509.

[15] Sen, P., & Puranam, P. (2022). Do alliance portfolios encourage or impede new business practice adoption? Theory and evidence from the private equity industry. Strategic Management Journal, 43(11), 2279–2312.

[16] Cao, Y., Ou, C., Dai, H., & Liu, Y. (2023). Research on the impact of network characteristics of disruptive innovation R&D alliance on enterprise innovative performance: The regulatory role of institutional environment. Journal of Jiangxi University of Finance and Economics, (01), 24–35.

[17] Liang, J., & Xie, E. (2020). Impact of divisive faultlines of multipartner R&D alliances on the firm innovation performance. Soft Science, 34(04), 26–30.

[18] Adelman, D. E., & Deangelis, K. L. (2006). Patent metrics: The mismeasure of innovation in the biotech patent debate. Texas Law Review, 85, 1677.

[19] Rothaermel, F. T., & Deeds, D. L. (2004). Exploration and exploitation alliances in biotechnology: A system of new product development. Strategic Management Journal, 25(3), 201–221.

[20] Wu, Q., Luo, X., Slotegraaf, R. J., & Aspara, J. (2015). Sleeping with competitors: The impact of NPD phases on stock market reactions to horizontal collaboration. Journal of the Academy of Marketing Science, 43(4), 490–511.

[21] Zheng, Y., & Yang, H. (2015). Does familiarity foster innovation? The impact of alliance partner repeatedness on breakthrough innovations. Journal of Management Studies, 52(2), 213–230.

[22] Colombo, M. G., Doganova, L., & Piva, E. (2015). Hybrid alliances and radical innovation: The performance implications of integrating exploration and exploitation. The Journal of Technology Transfer, 40, 696–722.

[23] Lin, Z., Yang, H., & Demirkan, I. (2007). The performance consequences of ambidexterity in strategic alliance formations: Empirical investigation and computational theorizing. Management Science, 53(10), 1645–1658.

[24] Li, W., & Wang, L. (2019). Strategic choices of exploration and exploitation alliances under market uncertainty. Management Decision, 57(11), 3112–3133.

[25] Li, Y., Jiao, W., & Yang, Z. (2023). Enterprise digital transformation and auditor risk decision-making. Finance and Accounting Monthly, 44(19), 111–119.

[26] Du, S., & Ma, L. (2024). The speed of digital transformation and debt financing costs: From the perspective of "cost reduction" and "deleveraging". Journal of Audit & Economics, 39(02), 52–62.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

To Be or Not to Be? The Influence of Ambidextrous Leadership Behaviour on Ambidextrous Innovation: A Moderated Mediation Model

Qian Gao*

School of Management, Xi'an Polytechnic University, Xi'an, Shaanxi, 710048, China

Corresponding author: Qian Gao, rewolfnuseliauk@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Based on the social information processing theory, this research constructed a process model of "ambidextrous leadership-goal orientation-ambidextrous innovation" and examined the moderating role of voice endorsement. An empirical test was conducted with the questionnaire data from 365 hospital staff. An empirical analysis of data from 365 hospital staff questionnaires revealed that: opening leadership behaviour positively affects employees' exploratory innovation, and learning goal orientation mediates the relationship; while closing leadership behaviour positively affects employees' exploitative innovation, and performance goal orientation mediates this relationship. Ambidextrous leadership positively influences exploratory innovation through learning goal orientation, and positively influences exploitative innovation through performance goal orientation. Voice endorsement positively moderates the relationship between opening leadership behaviour and learning goal orientation, and the relationship between closing leadership behaviour and performance goal orientation. Moreover, voice endorsement positively moderates the mediating role of learning goal orientation and performance goal orientation. This research provides insights to promote subordinates' ambidextrous innovation behaviour and improve organizational management.

Keywords: Ambidextrous Leadership; Goal Orientation; Ambidextrous Innovation; Voice Endorsement; Social Information Processing Theory

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1410

1. Introduction

Technological innovation in medicine is a crucial aspect for large comprehensive hospitals to lead the advancement of medical science and technology, and to promote significant progress in the field[1]. It is an inevitable requirement for the high-quality development of medical institutions. Medical personnel, as the key subjects of healthcare institutions, serve as the primary driving force for medical technological innovation, and the level of their innovation significantly determines the overall innovation capacity of these institutions[1]. To drive high-quality development in healthcare institutions through innovation, it is essential for medical professionals to not only acquire new knowledge, develop new skills, explore new opportunities, and experiment with new methods, known as exploratory innovation (exploration innovation)[2,3], but also to leverage existing knowledge to create value, enhance work efficiency, and optimize processes, referred to as exploitative innovation (exploitation innovation)[3,4]. Exploratory innovation and exploitative innovation constitute ambidextrous

innovation[5]. How to promote employees' ambidextrous innovative behaviour has become one of the current hot topics in organizational behaviour research. Leadership is recognized as one of the key factors that influence employee behaviour[6]. Research has found that transformational leadership and inclusive leadership can affect employees' ambidextrous innovation behaviours[7]. However, the conclusions are inconsistent[8,9]. According to the literature, transformational leadership negatively influence exploratory and exploitative employee innovation[10]. There are also studies that show transformational leadership facilitates employees' exploratory and exploitative innovation[9].

The reason for the inconsistent conclusions is that there is a contradiction between a singular leadership style and the necessity for ambidextrous innovation. Exploratory innovation and exploitative innovation entail different cognition, reasoning, and resources[5]. Exploratory innovation requires employees to go beyond existing knowledge, use diverse cognitive support for significant and radical innovation, and attempt to adopt new methods in a long-term period[11]. Pursuing exploratory innovation takes a significant investment of time and resources[12,13]. Exploitative innovation requires employees to employ existing knowledge in order to reinforce current technologies through a more consistent cognitive approach[14]. Exploitative innovation takes less time and resources for implementation since it adopts short-term goals and expands existing knowledge[12,13]. Therefore, promoting ambidextrous innovation among employees requires a leadership style that can flexibly flip between different approaches. However, existing research primarily explores the impact of single leadership behaviour on employees' ambidextrous innovation and falls short of addressing the tensions in antecedents of ambidextrous innovation[15]. Ambidextrous leadership is a paradoxical leadership model consisting of two complementary leadership behaviours and allows for flexible transitions between behavioural strategies[16]. It embodies integrated reasoning logic and can co-develop employees' explorative and exploitative behaviours[17]. However, literature on the relationship and mechanism between ambidextrous leadership and ambidextrous innovation remains notably limited[10].

The Social Information Processing Theory indicates that leaders are pivotal information sources for organizations[18]. They provide employees with not only general contextual information but also specific information through particular events[19]. Employees' attention, construction, and interpretation towards the information will have a differentiated impact on their attitudes and behaviours. This provides us with a new theoretical perspective to explore the relationship between ambidextrous leadership and employees' ambidextrous innovative behaviour. On the one hand, employees can identify and perceive general social contextual information clues conveyed by leaders with different leadership styles[20]. These contextual clues can affect employees' behaviours by influencing their attitudes, motivation, and goal[21]. The goal orientation of employees, shaped by contextual cues in the work environment, is dynamic and adaptable[22]. Employees' goal orientation can be affected by ambidextrous leadership and become antecedents of employees' innovative behaviour[23,24]. Different goal orientations exert distinct effects on different innovative behaviours[25]. Hence, this study introduces goal orientation as a mediator to investigate the effects of learning goal orientation and performance goal orientation in mediating the relationship between ambidextrous leadership and ambidextrous innovation. On the other hand, employees exhibit cognitive responses to specific information conveyed through particular events[19,21]. Such information can influence the employee's attention and information processing, thus shaping their goal orientation and behaviour[21].

Meanwhile, during the process of cultivating medical staff with individual innovative behaviors, frequent suggestions from these professionals enable leaders to respond swiftly and accurately to changes in the external environment, benefiting the long-term development of the industry[26]. However, what pivotal role does voice endorsement, as an essential managerial response, play in the relationship between ambidextrous leadership and individual ambidextrous innovative behaviors? Voice endorsement indicates leaders paying attention to, supporting, or implementing employees' suggestions[27]. This enables employees to perceive and interpret leaders' attitudes towards innovation based on their responses to suggestions. As a result, not only can leadership style influence employee goal orientation, but so can responses from leaders in specific situations[28]. Therefore, this study investigates the boundary effect of voice endorsement as a response and source of information clues on the relationship between ambidextrous leadership and ambidextrous innovation.

Overall, based on social information processing theory, this study views ambidextrous leadership as a general source of social contextual cues, and we introduce learning goal orientation and performance goal orientation to construct a research

framework of "ambidextrous leadership - goal orientation - ambidextrous innovation", and we verify the mediating effect of goal orientation. Furthermore, this study investigates the moderated mediation effect of leaders' voice endorsement by incorporating the perspective of specific event-related information cues from ambidextrous leadership, thereby enriching the research perspective and theoretical framework of ambidextrous leadership and ambidextrous innovation. This study also provides insights for organizations on promoting ambidextrous innovation behaviour and organizational innovation capability.

2. Theory and Hypotheses

2.1 Ambidextrous Leadership and Ambidextrous Innovation

Ambidextrous leadership is an approach that integrates dualism theory and leadership theory, giving rise to a novel leadership behaviour that combines two complementary leadership styles [29,30]. It embodies the cognitive perspective of 'both/and', and it involves the integration and coordination of contradictory activities, with the proposition of leaders' transition between two distinct leadership behaviours according to different situations[29,31,32]. Scholars primarily investigate ambidextrous leadership behaviour from four perspectives, the cognitive perspective (opening leadership and closing leadership)[33], the convention perspective (transformational leadership and transactional leadership)[32], the power perspective (empowering leadership and directive leadership)[34], and the benevolence and authority perspective (benevolent leadership and authoritative leadership) [35]. Excellent leaders demonstrate cognitive complexity, enabling them to shift cognitive modes when confronting complex environments[17]. This is crucial for exerting leadership strategies on specific contextual demands. Therefore, we investigate the relationship between ambidextrous leadership and ambidextrous innovation from a cognitive perspective of ambidextrous leadership.

Opening leadership behaviour refers to a leadership style that enhances heterogeneous employee behaviours by encouraging them to engage in different activities, explore new approaches, and support them in challenging the status quo[33,36]; closing leadership behaviour indicates a leadership style that reduces employee behavioural diversity through corrective measures, specific guidance, and goal supervision[33]. The social information processing theory holds that leaders are information sources for employees, and employees shape their cognition and behaviour in response to leaders' behaviour[37]. On the one hand, opening leadership promote employees' innovation willingness and information processing abilities by encouraging independent thinking and behaviour and trial-and-error[33]. This is conducive to promoting exploratory innovation. The literature indicates that opening leadership encourages employees to view the constantly changing environment as an opportunity and helps create conditions to support explorative innovation[38]. On the other hand, closing leadership signals employees that they are expected to follow established routines, ensuring that their behaviour aligns with organizational conventions[10,38]. Therefore, under closing leadership behaviour, employees tend to focus on specific goals and responsibilities and prefer a stable and predictable work environment, which promotes exploitative innovation[39,40]. Thus:

Hypothesis 1a: Opening leadership behaviour has a positive effect on employees' exploratory innovation.

Hypothesis 1b: Closing leadership behaviour has a positive effect on employees' exploitative innovation.

2.2 The Mediating Role of Goal Orientation

Goal orientation refers to an individual's cognitive and behavioural pattern of pursuing goals [41,42], including learning goal orientation and performance goal orientation. Learning goal orientation refers to setting goals to enhance their understanding and comprehension of new concepts[43]. Performance goal orientation pertains to individuals striving to receive favorable evaluations of their abilities or to avoid negative evaluations[42,43]. Individual goal orientation is dynamic and can be affected by external information, which leads to behavioural changes[3]. Therefore, integrating with the social information processing theory, this study investigates the mechanism of ambidextrous leadership on ambidextrous innovation from a goal-orientation perspective.

The social information processing theory indicates that employees' cognition and attitude are affected by assessing, processing, and interpreting general social contextual information in the work environment[21], leading to different goal orientations[44]. Opening leadership encourages employees to actively explore new methods and learn from mistakes and failures. This promotes independent and critical thinking in employees and helps deepen their understanding and

implementation of knowledge to gain further work experience[6]. Moreover, opening leadership motivates employees to explore and overcome barriers and thus promotes employees' learning goal orientation[23,38]. Literature indicates that opening leadership behaviour helps establish a climate of knowledge and information sharing in an organization, which can further nurture learning goal orientation[45]. In contrast, closing leadership emphasizes following plans and controlling work progress through routines, which can improve employees' efficiency[23,38]. Employees are more inclined to prove that they can complete tasks more quickly and effectively, thus forming a high-performance goal orientation. Moreover, closing leadership provides employees with norms and standards, requiring employees to comply with routines and transmitting signals of the significance of performance to employees[46]. This will encourage employees to improve performance by reducing behavioural differences, thereby enhancing performance goal orientation.

As can be seen, ambidextrous leadership can effectively respond to complex and dynamic situations, motivating employees. Goal orientation, as a type of achievement motivation, is a significant predictor of individual behaviour and helps explain the differences in exploratory and exploitative behaviour in the workplace[6]. The theory of social information processing theory also indicates that employees adjust their behaviour by interpreting different information[21]. Employees with high learning goal orientation believe their abilities can be improved through efforts, so they focus on developing new skills and acquiring knowledge, fostering strong motivation[47]. As a result, these employees are inclined to exhibit autonomous behaviours[24], such as spontaneously engaging in exploratory innovation. Learning goal orientation enables individuals to embrace challenges as opportunities rather than threats[42]. Therefore, employees with high learning goal orientations are willing to implement exploratory innovation with risks and uncertainties to achieve themselves. Employees with performance goal orientation focus on demonstrating their abilities and gaining recognition from others[47]. These employees focus more on improving existing work to enhance performance and less on individual capacity development[48]. Therefore, employees with performance goal orientation tend to do exploitative innovation rather than exploratory innovation.

In summary, opening leadership encourages experimentation and breakthroughs, which promotes learning activities and exploratory innovation in employees. Closing leadership prefers efficiency and standardisation, which motivates performance-oriented and exploitative innovation in employees. Thus:

Hypothesis 2a: The employee learning goal orientation mediates the relationship between opening leadership behaviour and employees' exploratory innovation.

Hypothesis 2b: The employee performance goal orientation mediates the relationship between closing leadership behaviour and employees' exploitative innovation.

The ambidextrous leadership reflects the coordination between the contradictions and tensions of different leadership behaviours, i.e. the "both/and" thinking logic rather than the "either/or" logic[49,50]. Ambidextrous leadership encourages employees to constantly try new things and challenge the status quo through opening behaviour, stimulating their creative enthusiasm and innovative behaviour[6]. Ambidextrous leadership can also alleviate the suppression of divergent thinking caused by closing behaviour[51]. In addition, closing behaviour establishes norms and routines for employees to reduce the coordination problems raised by opening behaviour[23]. Therefore, ambidextrous leadership generates synergies through flexibly switching between two complementary leadership behaviours. Employees are encouraged not only to adopt a learning goal orientation, fostering exploratory innovation but also to adopt a performance goal orientation, promoting exploitative innovation. Thus:

Hypothesis 3a: The employee learning goal orientation plays a mediating role between ambidextrous leadership behaviour and employees' exploratory innovation.

Hypothesis 3b: The employee performance goal orientation plays a mediating role between ambidextrous leadership behaviour and employees' exploitative innovation.

2.3 The Moderating Effect of Voice Endorsement

Voice endorsement is the extent to which leaders value and recognize suggestions made by employees to improve the organization[52]. Voice endorsement provides employees with opportunities and supportive resources from leaders to develop the organization through suggestions[53,54]. Hunton et al. (1996) found that compared with voice endorsement,

voice rejection results in a decrease of 41% in employees' output[6,55]. This indicates that voice endorsement is significant in leaders influencing employees' work. According to the social information processing theory, leadership behaviour and voice endorsement are both crucial information clues in the workplace environment. Leadership behaviour is a general social contextual information clue directed towards employees[19]. Voice endorsement is specific information based on the feedback from employees and leaders regarding specific advice events[56]. Employees' goal orientation and information processing may be affected when they perceive particular information[41]. Therefore, this study examines the moderating role of voice endorsement as a special information clue in the relationship between leadership behaviour and goal orientation.

As mentioned, opening leadership positively affect employees' learning goal orientation, as employees focus on ability enhancement due to the leader's encouragement. Research indicates that opening leadership inspire employees to generate innovative thoughts and initiatives[6]. Employees who make learning and explorative suggestions are more likely to be noticed and recognized by opening leaders. Higher voice endorsement enhances employees' perception of leadership encouragement for learning and exploration, thereby strengthening employees' learning goals. Moreover, voice endorsement serves as positive feedback to make employees perceive appreciation and trust in leaders[53,54,57]. This, in turn, strengthens employees' learning motivation and improves their learning goal orientation. Employees with lower voice endorsement may feel unvalued by leaders, which may trigger negative reactions such as a weakened learning goal orientation[57].

Correspondingly, closing leadership positively affect employees' performance goal orientation by emphasizing task completion. Research has shown that closing leadership focuses on responsibility awareness and facilitates employees to provide preventive suggestions on issues that impede productivity[6]. Therefore, closing leaders prioritize performance improvement and utilization suggestions proposed by employees. Higher voice endorsement reflects the recognition and appreciation of closing leadership towards the ideas and suggestions of employees[53]. This empowers employees to focus more on responsibilities and tasks and enhances their performance goal orientation[6]. Lower voice endorsement leads to negative self-assessments, causing employees to doubt their abilities to accomplish tasks or solve problems and weakening performance goal orientation[58]. Thus:

Hypothesis 4a: Voice endorsement positively moderates the relationship between opening leadership behaviour and employees' learning goal orientation.

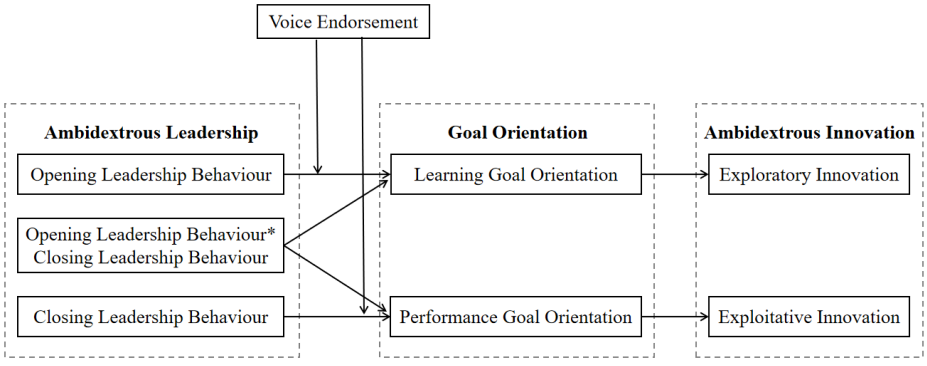
Hypothesis 4b: Voice endorsement positively moderates the relationship between closing leadership behaviour and employees' performance goal orientation.

In conclusion, under higher voice endorsement, the signal of opening leadership explicitly encourages employees to learn and explore, which enhances employees' learning goal orientation and promotes exploration; closed leadership signals a straightforward emphasis on efficiency and routines, which enhances the performance goal orientation and exploitative innovation in employees. Thus:

Hypothesis 5a: Voice endorsement positively moderates the mediating role of employees' learning goal orientation between opening leadership behaviour and employees' exploratory innovation.

Hypothesis 5b: Voice endorsement positively moderates the mediating role of employees' performance goal orientation between closing leadership behaviour and employees' exploitative innovation.

Figure 1. Presents the conceptual model.



3. Method

3.1 Sample and Procedure

We collect data from doctors, medical technicians, and nurses in hospitals in Sichuan, Chongqing, Shaanxi, and other areas in China. Individuals in medical technology services and health management specialize in technology R&D and service improvement, and they are more likely to take exploration and exploitation innovation activities. Literature has studied the impact of different leadership styles on individual innovation behaviour among the healthcare community in public hospitals[59]. Therefore, focusing on the medical staff and investigating their ambidextrous innovation is consistent with this study's theoretical requirements. We first entered the questionnaire items into the Credamo platform and generated the electronic link, then we communicated with and asked the target hospital's managers to assist us in sending the questionnaire to their staff. We collected 388 questionnaires, and we got 365 valid questionnaires with a 93.83% response rate after data washing (i.e., we washed the questionnaires that were filled out for a particularly short time and those that were filled with patterns to guarantee data quality). Of these valid questionnaires, 28.22% were male and 71.78% were female; regarding age, 22-31-year-olds accounted for 31.78%, while 32-41-year-olds accounted for 27.12%, accounting for more than half of all respondents; those with a bachelor's degree accounted for the largest proportion of respondents (60.55%); 50.41% of respondents had worked for more than ten years, 18.90% for 6-10 years, and 30.69% had worked for less than or equal to five years.

3.2 Measures

The questionnaire was developed using established scales that have been widely used in existing research. A standard back-translation procedure was performed to ensure not only the accuracy of the translation but also the adaptation of items to the Chinese context. The scales were measured on a five-point Likert scale, with "1" representing "strongly disagree" and "5" representing "strongly agree".

Ambidextrous Leadership. Ambidextrous leadership is the product of opening and closing leadership[6]. Opening and closing leadership are measured by Rosing et al.(2011)[60]. The opening leadership behaviour scale has five items, such as "My immediate superior allows me to present different ideas," while the closing leadership behaviour scale consists of five items, such as "My immediate superior establishes standards for completing tasks".

Goal Orientation. The measurement of learning goal orientation and performance goal orientation originates from Vandewalle(1997)[61]. An example of an item of learning goal orientation is "I am willing to choose challenging tasks so that I can learn a lot from them". An example of an item of performance goal orientation is "I work hard to find ways to prove my abilities to others in my job".

Ambidextrous Innovation. Ambidextrous innovation is measured by exploratory innovation and exploitative innovation based on the study of Mom et al.(2009)[62]. We use five items to measure the exploratory innovation. An example of an item is "In my work, the innovation I make requires seeking new opportunities in technology/process/service"; We use four items to measure the exploitative innovation. An example of an item is "In my work, the innovation I make can be completed using existing knowledge".

Voice Endorsement. Following the research of Burris (2012)[63],we measure voice endorsemnt using a five-item scale. An example of an item is "My immediate supervisor considers my job suggestions or opinions valuable".

Control Variables. The control variables in this study include organizational size, job position, gender, and age.

4. Results

4.1 Reliability and Validity Test

We conducted several tests to ensure the reliability and validity of the data. Regarding reliability, the Cronbach's alpha values for opening leadership (0.915), closing leadership (0.870), learning goal orientation (0.894), performance goal orientation (0.852), exploratory innovation (0.935), exploitative innovation (0.839), and voice endorsement (0.910) were analysed using SPSS 24.0. The results show that the Cronbach's alpha values of all constructs were greater than 0.7, which exhibits good reliability.

Regarding validity, we conducted the KMO value and Bartlett's sphericity test in SPSS 24.0 and tested the convergent validity and discriminant validity by AMOS 24.0. Convergent validity results show that the factor loadings of all the items were greater than 0.6. The average extracted variances (AVE) of the constructs were 0.686 (opening leadership), 0.578 (closing leadership), 0.693 (learning goal orientation), 0.600 (performance goal orientation), 0.786 (exploratory innovation), 0.637 (exploitative innovation), and 0.673 (voice endorsement), the AVEs were above 0.5, indicating a high discriminant validity. The constructs' composite reliabilities (CR) were 0.916, 0.872, 0.900, 0.856, 0.936, 0.840, and 0.911, all greater than 0.7, indicating a good convergent validity. Table 1 presents the confirmatory factor analysis, which indicates that the seven-factor model is best-fitting and our constructs have good discriminant validity.

Table 1. Results of the Confirmatory Factor Analysis.

Model	χ^2	df	χ^2/df	RMSEA	RMR	CFI	TLI	IFI
seven-factor	1005.579	384	2.619	0.067	0.038	0.927	0.917	0.927
six-factor	1129.776	390	2.897	0.072	0.041	0.913	0.903	0.913
five-factor	1519.835	395	3.848	0.088	0.058	0.868	0.854	0.868
four-factor	1681.229	399	4.214	0.094	0.06	0.849	0.836	0.850
three-factor	2628.022	402	6.537	0.123	0.09	0.738	0.717	0.739
two-factor	3074.075	404	7.609	0.135	0.096	0.686	0.662	0.687
single-factor	3844.130	405	9.492	0.153	0.101	0.596	0.566	0.597

Note: N=365. The seven-factor model is a hypothetical model; the six-factor model combines exploratory innovation and exploitative innovation into one factor; the five-factor model combines learning goal orientation and performance goal orientation into one factor; the four-factor model combines opening leadership and closing leadership into one factor; the three-factor model combines exploratory innovation, exploitative innovation, and voice endorsement into one factor; the two-factor model combines exploratory innovation, exploitative innovation, voice endorsement, learning goal orientation, and performance goal orientation into one factor; and the single-factor model combines all seven variables into one factor.

4.2 Common Method Biases

This study employed a method of controlling for unmeasured latent method factors to examine the common method biases. We added a common method factor to the seven-factor model (Table 1) to make an eight-factor comparative model. We compared the fit indices of the two models, and the changes in each index were below 0.03 (△ CFI=0.014, △ TLI=0.010, △ ILI=0.014; △ RMSEA=0.004, △ RMR=0.006), indicating that adding a common method factor did not improve the model. Therefore, the common method variance is not a concern in our study.

4.3 Descriptive Statistics and Correlation Analysis

Table 2 presents the descriptive statistical analysis. As shown in Table 2, opening leadership demonstrates a positive correlation with learning goal orientation ($r=0.521$, $p<0.01$), while closing leadership shows a positive correlation with performance goal orientation ($r=0.353$, $p<0.01$). Additionally, learning goal orientation is positively correlated with exploratory innovation ($r=0.654$, $p<0.01$), and performance goal orientation is positively correlated with exploitative innovation ($r=0.473$, $p<0.01$). These findings provide preliminary support for the research.

Table 2. Descriptive statistical analysis.

Variables Results of the Confirmatory Factor Analysis	1	2	3	4	5	6	7	8	9	10	11
1. Size of organization	1										
2. Job positions	0.021	1									
3. Gender	0.059	-0.105*	1								
4. Age	-0.114*	-0.001	-0.202**	1							
5. Opening leadership behaviour	-0.043	-0.022	0.069	0.149**	1						
6. Closing leadership behaviour	-0.033	0.023	0.055	0.096	0.727**	1					
7. Learning goal orientation	-0.081	0.106*	-0.050	0.082	0.521**	0.481**	1				
8. Performance goal orientation	-0.115*	0.030	-0.045	0.130*	0.323**	0.353**	0.527**	1			
9. Exploratory innovation	-0.078	0.210**	0.024	0.033	0.555**	0.512**	0.654**	0.443**	1		
10. Exploitative innovation	-0.063	0.110*	0.023	0.058	0.484**	0.502**	0.563**	0.473**	0.750**	1	
11. Voice endorsement	-0.077	0.051	-0.015	0.165**	0.704**	0.607**	0.468**	0.260**	0.505**	0.423**	1
Mean	3.230	2.403	1.718	3.093	3.993	4.031	4.047	3.702	3.956	3.972	3.815
S D	1.009	1.305	0.451	0.976	0.804	0.719	0.785	0.908	0.806	0.761	0.752

Note: N = 365; * p < 0.05; ** p < 0.01; *** p < 0.001 (two-tailed), the same below.

4.4 Test of Main Effect and Mediation Effect

Figure 2 shows the main and mediation effects testing by AMOS 24.0 (model fit: 2/df=2.301, RMSEA=0.060, CFI=0.923, TLI=0.914, IFI=0.923). The results indicate that opening leadership behaviour is significantly positively correlated with exploratory innovation (b=0.428, p<0.001), and closing leadership behaviour is significantly positively correlated with exploitative innovation (b=0.489, p<0.001). Thus, Hypothesis 1a and Hypothesis 1b are supported. Furthermore, there is a significant positive correlation between open leadership and learning goal orientation (b=0.661, p<0.001), and a significant positive correlation between closed leadership and performance goal orientation (b=0.547, p<0.001). Additionally, learning goal orientation is significantly positively correlated with exploratory innovation (b=0.353, p<0.001), and performance goal orientation is significantly positively correlated with exploitative innovation (b=0.309, p<0.001).

Figure 2. Path analysis result.

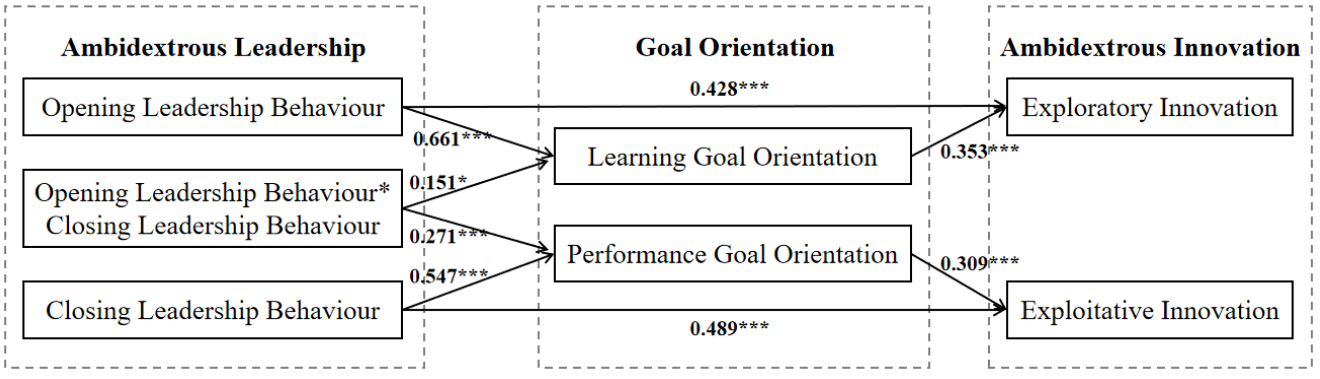


Table 3 presents the results of the indirect effect analysis. We adopted the Bootstrapping method (Table 3) with 5000 Bootstrap sample sizes and a 95% confidence level for the confidence interval. The value between BootLLCI and BootULCI excluding 0 indicates that the mediation effect is significant. Opening leadership positively influences exploratory innovation through goal orientation learning with an indirect effect value of 0.251 and a 95% confidence interval of [0.159, 0.388], supporting Hypothesis 2a. Closing leadership positively influences exploitative innovation through performance goal orientation with an indirect effect value of 0.176 and a 95% confidence interval of [0.093, 0.300], supporting Hypothesis 2b. Ambidextrous leadership exerts a positive impact on exploratory innovation through learning goal orientation, with an indirect effect value of 0.051 and a 95% confidence interval of [0.007, 0.805], supporting Hypothesis 3a. Ambidextrous leadership also positively influences exploitative innovation through performance goal orientation with an indirect effect value of 0.068 and a 95% confidence interval of [0.026, 0.507], supporting Hypothesis 3b.

Table 3. Results of the mediation effects.

Path	Estimate	SE	[BootLLCI, BootULCI]
Opening leadership behaviour→Learning goal orientation→-Exploratory innovation	0.251	0.064	[0.159,0.388]
Closing Leadership behaviour→Performance goal orientation→Exploitative innovation	0.176	0.055	[0.093,0.300]
Ambidextrous leadership→Learning goal orientation→Exploratory innovation	0.051	0.284	[0.007,0.805]
Ambidextrous leadership→Performance goal orientation→-Exploitative innovation	0.068	0.142	[0.026,0.507]

Bootstrapping sample size is 5000; 95% confidence intervals presented

4.5 Test of Moderating Effects and Moderated Mediating Effects

This study used PROCESS v3.3 to test the moderated mediation effect of voice endorsement with the Bootstrapping technique (Table 4). Firstly, voice endorsement positively moderates the relationship between opening leadership and learning goal orientation, and the results indicate a significant indirect effect ($b=0.101$, $p<0.05$) with a 95% confidence interval of [0.012,0.190]. Thus, Hypothesis 4a is supported, as shown in Fig. 3. Secondly, voice endorsement positively moderates the relationship between opening leadership and learning goal orientation ($b=0.213$, $p<0.01$) with a 95% confidence interval of [0.093,0.333], supporting Hypothesis 4b, as shown in Fig. 4.

Table 4. Results of the moderating effects and moderating mediating effects.

Path	Effect	Standard error	95% CI	
			Lower limit	Upper limit
Opening leadership behaviour →Learning goal orientation	0.101	0.045	0.012	0.190
Closing Leadership behaviour →Performance goal orientation	0.213	0.061	0.093	0.333
Opening leadership behaviour →Learning goal orientation →Exploratory innovation	0.050	0.031	0.001	0.120
Closing Leadership behaviour →Performance goal orientation →Exploitative innovation	0.061	0.021	0.020	0.102

Bootstrapping sample size is 5000; 95% confidence intervals presented.

Figure 3. The moderating effect of voice endorsement between opening leadership behaviour and learning goal orientation.

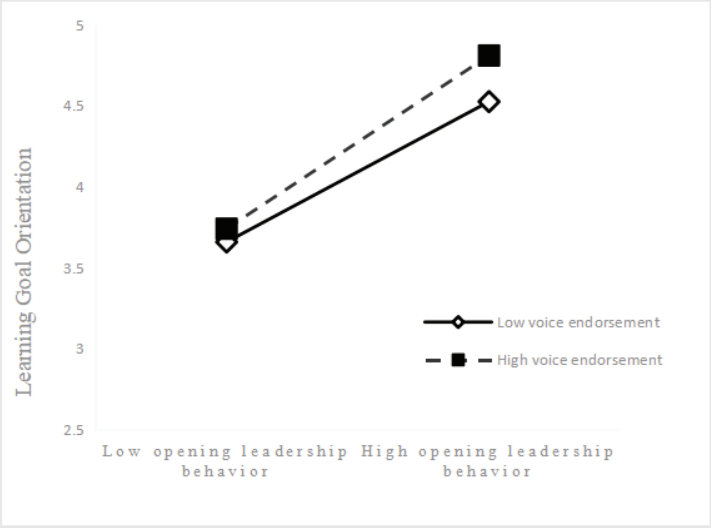
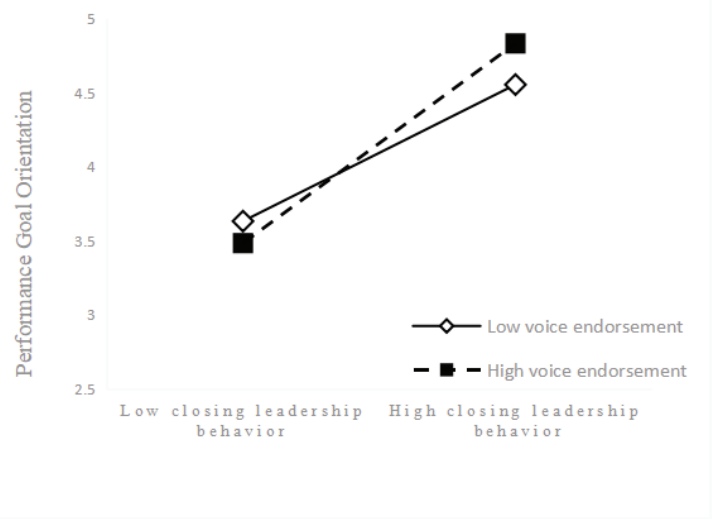


Figure 4. The moderating effect of voice endorsement between closing leadership behaviour and performance goal orientation.



Moreover, we applied PROCESS v3.3 to investigate the moderated mediating effect of voice endorsement on goal orientation between ambidextrous leadership and ambidextrous innovation (Table 4). The mediating role of learning goal orientation between opening leadership and exploratory innovation is positively moderated by voice endorsement, with an indirect effect value of 0.050 and a 95% confidence interval of [0.001, 0.120]. Thus, Hypothesis 5a is supported. In addition, voice endorsement positively moderates the mediating role of performance goal orientation between closing leadership and exploitative innovation, with an indirect effect value of 0.061 and a 95% confidence interval of [0.020, 0.102]. Thus, Hypothesis 5b is supported.

5. Discussion

5.1 Conclusion

Based on the social information processing theory, this study examines the mechanism of ambidextrous leadership on ambidextrous innovation from the goal orientation perspective. The results indicate that: (1) Opening leadership behaviour positively influences exploratory innovation, whereas closed leadership behaviour positively influences exploitative innovation. (2) Learning goal orientation partially mediates the relationship between opening leadership behaviour and exploratory innovation, whereas performance goal orientation partially mediates the relationship between closing leadership behaviour and exploitative innovation. Moreover, ambidextrous leadership, resulting from the product of open and closed leadership, has a positive impact on exploratory innovation through learning goal orientation and a positive influence on

exploitative innovation through performance goal orientation. (3) This study examines the positive moderated effect of voice endorsement on the indirect relationship between leadership behaviour and ambidextrous innovation through goal orientation.

5.2 Theoretical Contributions

Our study makes three contributions to existing research on this topic. First, based on the social information processing theory, this study reveals the mechanism of ambidextrous leadership on employees' ambidextrous innovation, which expands the theoretical research perspective of ambidextrous leadership. Prior studies have primarily focused on the effect of ambidextrous leadership behaviour on team innovation, team creativity, and employee innovative behaviour from the perspectives of emotional cognitive theory[6], interactive cognitive theory[17], dual-factor theory[23,64], and social cognitive theory[64]. However, most of these studies have explored the impact of ambidextrous leadership behaviour as a general social contextual cue on outcomes. This study integrates ambidextrous leadership and voice endorsement into a comprehensive information framework from the perspective of social information processing. We analyze the role of general social contextual cues conveyed by leadership behaviour and elucidate the role of specific informational cues from leaders' voice endorsement. Our study thus not only provides theoretical support for a comprehensive process of ambidextrous innovation behaviour but also supplements the research of the social information processing theory.

Second, this study reveals the influence and mechanism of ambidextrous leadership on ambidextrous innovation. The findings not only furnish empirical evidence for investigating the antecedents of ambidextrous innovation but also establish a connection between voice endorsement and goal orientation within ambidextrous leadership and innovation research. Existing research has predominantly examined the relationship between ambidextrous leadership and innovation through the perspectives of team interaction and dyadic learning[6,51]. This study explores how ambidextrous leadership influences employees' ambidextrous innovation, examining the mediating effects of learning and performance goal orientations. Our study introduces a novel perspective to examine the impact of ambidextrous leadership on ambidextrous innovation, contributing to a comprehensive understanding of its underlying processes.

Third, this study reveals the moderating role of voice endorsement in the relationship between ambidextrous leadership and ambidextrous innovation. This advances research on the boundary conditions governing how ambidextrous leadership influences ambidextrous innovation. While existing literature predominantly focuses on the factors influencing or the outcomes of voice endorsement, this study considers leaders' voice endorsement as a specific form of social information. We incorporate voice endorsement into the research framework of "ambidextrous leadership - goal orientation - ambidextrous innovation" based on the social information processing theory. By revealing the moderating and moderated mediating effect of voice endorsement, we supplement the boundary research on the relationship between ambidextrous leadership and ambidextrous innovation. Moreover, it extends the research on voice endorsement.

5.3 Practical Contributions

Our study offers several significant practical implications, helping illuminate how leaders achieve employees' ambidextrous innovation. First, managers can nurture and develop ambidextrous thinking and leadership abilities to stimulate employees to participate in ambidextrous innovation activities. Ambidextrous leadership thinking allows for the flexible handling of contradictions that may arise during the innovation process. Therefore, leaders should attempt to switch between opening and closing leadership styles in different situations to achieve the synergy effect of ambidextrous leadership.

Second, managers can help employees develop different goal orientations to cultivate ambidextrous innovation behaviours. Learning goal orientation can motivate employees to gain new knowledge and abilities. Performance goal orientation can improve employees' performance and efficiency. For example, managers can encourage employees to try new approaches to complete tasks or inspire employees' work enthusiasm through role division and task allocation. Moreover, managers can offer rewards to stimulate employees' achievement motivation.

Third, managers should consider valuable suggestions from employees. Organizations cannot thrive without members' voices; therefore, managers should encourage employees to offer suggestions and refine the mechanism for endorsing those suggestions. Moreover, it is critical to acknowledge employees' voices and provide feedback, which promotes an awareness of recognition and value in employees, strengthening their connection with leaders and the organization.

5.4 Limitations and Future Directions

Despite its contributions, this study also has limitations. First, this study used the cross-sectional design of questionnaires, which limits our ability to assess causality relationships. Future research can further examine the relationship through longitudinal data. Second, this study used a multiplication of opening and closing leadership behaviour to represent ambidextrous leadership. Although this is a general measurement of "ambidextrous" leadership[6,51], there have also been studies using "difference" or "sum" to measure "ambidextrous" leadership. Future research can try alternative scales for ambidextrous leadership and validate our conclusions. Third, this study revealed the boundary effects of voice endorsement, while not distinguishing the contents of voice endorsement. Previous research has found that leaders react differently to different types of suggestion content[63]. Therefore, future research can further explore the effect of different voice endorsement content on the relationship between ambidextrous leadership behaviours and goal orientation.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Mutonyi, B. R., Slåtten, T., & Lien, G. (2021). Fostering innovative behavior in health organizations: A pls-sem analysis of norwegian hospital employees. BMC Health Services Research, 21, 470.

[2] Deng, X., Ali, H., & Aboelkheir, H. (2025). Unleashing creative potential: The role of stable institutional ownership in exploratory innovation. Research Policy, 54, 105166.

[3] Jafari-Sadeghi, V., Garcia-Perez, A., Candelo, E., & Couturier, J. (2021). Exploring the impact of digital transformation on technology entrepreneurship and technological market expansion: The role of technology readiness, exploration and exploitation. Journal of Business Research, 124, 100–111.

[4] Sun, B., Zhang, Y., Zhao, Y., Mao, H., Kang, M., & Liang, T. (2025). Does continuous innovation failure lead firm innovation to hesitate to press forward?: Evidence from Chinese-listed technology-intensive industries firms. Journal of Business Research, 186, 114986.

[5] Cao, T., Li, W., Zhang, Y., & Zheng, M. (2024). Win–win or Lose–Lose? Dual logic, board factional faultlines, and ambidextrous innovation in state-owned enterprises. Technovation, 137, 103099.

[6] Zacher, H., & Rosing, K. (2015). Ambidextrous leadership and team innovation. Leadership & Organization Development Journal, 36, 54–68.

[7] Hoang, G., Luu, T. T., Nguyen, T. T., Du, T., & Le, L. P. (2022). Examining the effect of entrepreneurial leadership on employees' innovative behavior in SME hotels: A mediated moderation model. International Journal of Hospitality Management, 102, 103142.

[8] Gong, L., Liu, Z., Rong, Y., & Fu, L. (2021). Inclusive leadership, ambidextrous innovation and organizational performance: The moderating role of environment uncertainty. Leadership & Organization Development Journal, 42, 783–801.

[9] Stein, M., Schümann, M., & Vincent-Höper, S. (2021). A conservation of resources view of the relationship between transformational leadership and emotional exhaustion: The role of extra effort and psychological detachment. Work & Stress, 35, 241–261.

[10] Busola Oluwafemi, T., Mitchelmore, S., & Nikolopoulos, K. (2020). Leading innovation: Empirical evidence for ambidextrous leadership from UK high-tech SMEs. Journal of Business Research, 119, 195–208.

[11] Zhang, X., Sun, Y., & Zhang, Y. (2024). International R&D cooperation and Chinese firms' exploitative and exploratory innovation in a bifurcated world. Asia Pacific Journal of Management, 1–28.

[12] Hunter, S. T., Cushenbery, L. D., & Jayne, B. (2017). Why dual leaders will drive innovation: Resolving the exploration and exploitation dilemma with a conservation of resources solution. Journal of Organizational Behavior, 38, 1183–1195.

[13] Zhou, J., & Hoever, I. J. (2014). Research on workplace creativity: A review and redirection. Annual Review of Organizational Psychology and Organizational Behavior, 1, 333–359.

[14] Choi, J., Kim, Y.-K., & Roh, T. (2024). Unpacking the link between entrepreneurial proactiveness and exploitative innovation strategy: The role of brokerage position and open innovation. Technovation, 136, 103068.

[15] Agarwal, P., & Farndale, E. (2017). High-performance work systems and creativity implementation: The role of psychological capital and psychological safety. Human Resource Management Journal, 27, 440–458.

[16] Ajmal, M., Sareet, Z., & Islam, A. (2025). Unleashing innovation through employee voice behavior in the hotel industry: The impact of ambidextrous leadership on innovative work behavior. Journal of Hospitality Tourism and International, 8, 448–471.

[17] Liu, S., Hu, J., Li, Y., Wang, Z., & Lin, X. (2014). Examining the cross-level relationship between shared leadership and learning in teams: Evidence from China. The Leadership Quarterly, 25, 282–295.

[18] Li, Y. (2014). A multi-level model of individual information privacy beliefs. Electronic Commerce Research and Applications, 13, 32–44.

[19] Kietzmann, T. C., & König, P. (2015). Effects of contextual information and stimulus ambiguity on overt visual sampling behavior. Vision Research, 110, 76–86.

[20] Zheng, Y., Gao, Y. L., Li, M., & Dang, N. (2023). Leadership styles and employee pro-environmental behavior in the tourism and hospitality industry: A cognitive-affective personality system perspective. International Journal of Hospitality Management, 113, 103509.

[21] Salancik, G. R., & Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. Administrative Science Quarterly, 23, 224.

[22] Dust, S. B., Merchant, S. M., Wang, P., & Yao, X. (2025). The moderating role of narcissism on the impact of leader learning goal orientation on follower adaptivity through follower goal orientation: A social cognitive theory perspective. The Journal of Social Psychology.

[23] Biemann, T., Kearney, E., & Marggraf, K. (2015). Empowering leadership and managers' career perceptions: Examining effects at both the individual and the team level. The Leadership Quarterly, 26, 775–789.

[24] Rhee, Y. W., & Choi, J. N. (2017). Knowledge management behavior and individual creativity: Goal orientations as antecedents and in-group social status as moderating contingency. Journal of Organizational Behavior, 38, 813–832.

[25] Zia, N. U. (2020). Knowledge-oriented leadership, knowledge management behaviour and innovation performance in project-based SMEs. The moderating role of goal orientations. Journal of Knowledge Management, 24, 1819–1839.

[26] Huai, M., Wen, X., Liu, Z., Wang, X., Li, W., & Wang, M. (2024). Does voice endorsement by supervisors enhance or constrain voicer's personal initiative? Countervailing effects via feeling pride and feeling envied. Journal of Applied Psychology, 109, 1408–1430.

[27] Satterstrom, P., Kerrissey, M., & DiBenigno, J. (2021). The voice cultivation process: How team members can help upward voice live on to implementation. Administrative Science Quarterly, 66, 380–425.

[28] Li, L., Huang, G., & Yan, Y. (2022). Coaching leadership and employees' deviant innovation behavior: Mediation and chain mediation of interactional justice and organizational identification. Psychology Research and Behavior Management, 15, 3861–3874.

[29] Gianzina, O., & Paroutis, S. (2025). Ambidextrous leadership: A systematic review and roadmap for future research. Journal of Business Research, 189, 115151.

[30] Hansen, N. K., Güttel, W. H., & Swart, J. (2019). HRM in dynamic environments: Exploitative, exploratory, and ambidextrous HR architectures. The International Journal of Human Resource Management, 30, 648–679.

[31] Rosing, K., Frese, M., & Bausch, A. (2011). Explaining the heterogeneity of the leadership-innovation relationship: Ambidextrous leadership. The Leadership Quarterly, 22, 956–974.

[32] Schreuders, J., & Legesse, A. (2012). Organizational ambidexterity: How small technology firms balance innovation and support. Technology Innovation Management Review.

[33] Mascareño, J., Rietzschel, E. F., & Wisse, B. (2021). Ambidextrous leadership: Opening and closing leader behaviours to facilitate idea generation, idea promotion and idea realization. European Journal of Work and Organizational Psychology, 30, 530–540.

[34] Sagie, A., Zaidman, N., Amichai-Hamburger, Y., Te'eni, D., & Schwartz, D. G. (2002). An empirical assessment of the loose-tight leadership model: Quantitative and qualitative analyses. Journal of Organizational Behavior, 23, 303–320.

[35] Uhl-Bien, M., Riggio, R. E., Lowe, K. B., & Carsten, M. K. (2014). Followership theory: A review and research agenda. The Leadership Quarterly, 25, 83–104.

[36] Sonmez Cakir, F., & Adiguzel, Z. (2019). Evaluation of open leadership and innovation orientation on employees and culture of the organization. Business: Theory and Practice, 20, 432–445.

[37] Antunes, A., & Franco, M. (2016). How people in organizations make sense of responsible leadership practices: Multiple case studies. Leadership & Organization Development Journal, 37, 126–152.

[38] Kousina, E., Voudouris, I. (2023). The ambidextrous leadership-innovative work behavior relationship in the public sector: The mediating role of psychological ownership. Public Administration Review, 83, 1478–1495.

[39] Kebede, A. G., Terefe, S. D., & Ijigu, A. W. (2024). Ambidextrous leadership and academic staff innovative behavior at Debre Berhan University, Ethiopia: The mediating role of workplace happiness. Cogent Business & Management, 11, 2307564.

[40] Wang, S., Eva, N., Newman, A., & Zhou, H. (2021). A double-edged sword: The effects of ambidextrous leadership on follower innovative behaviors. Asia Pacific Journal of Management, 38, 1305–1326.

[41] Van Strydonck, I., Decramer, A., & Audenaert, M. (2025). Linking leaders' learning goal orientation to employees' innovative work behaviour, job satisfaction and emotional exhaustion: Leaders' provision of learning-oriented goals and feedback and coaching accuracy. European Journal of Work and Organizational Psychology, 34, 109–127.

[42] Vandewalle, D., Nerstad, C. G. L., & Dysvik, A. (2019). Goal orientation: A review of the miles traveled and the miles to go. Annual Review of Organizational Psychology and Organizational Behavior, 6, 115–144.

[43] Liu, M. (2024). Learning goal orientation, social capital, and entrepreneurial intentions: A multigroup analysis within entrepreneurship programs. The International Journal of Management Education, 22, 101043.

[44] Gao, S., Lin, X., & Lam, L. W. (2024). How learning-goal-oriented newcomers develop leader-member exchange (LMX): From agentic and communal behavior perspectives. Journal of Business Research, 176, 114569.

[45] Berraies, S., & Zine El Abidine, S. (2019). Do leadership styles promote ambidextrous innovation? Case of knowledge-intensive firms. Journal of Knowledge Management, 23, 836–859.

[46] Hu, W., Luo, J., Chen, Z., & Zhong, J. (2020). Ambidextrous leaders helping newcomers get on board: Achieving adjustment and proaction through distinct pathways. Journal of Business Research, 118, 406–414.

[47] Lin, Q., Lin, L., & Ye, D. (25). Factors influencing knowledge-sharing behaviors and learning effect: A multilevel investigation. Social Behavior and Personality, 43, 1683–1698.

[48] Malik, M. A. R., Choi, J. N., & Butt, A. N. (2019). Distinct effects of intrinsic motivation and extrinsic rewards on radical and incremental creativity: The moderating role of goal orientations. Journal of Organizational Behavior, 40, 1013–1026.

[49] Gebert, D., Boerner, S., & Kearney, E. (2010). Fostering team innovation: Why is it important to combine opposing action strategies? Organization Science, 21, 593–608.

[50] Trong Tuan, L. (2017). Reform in public organizations: The roles of ambidextrous leadership and moderating mechanisms. Public Management Review, 19, 518–541.

[51] Duc, L. A., Tho, N. D., Nakandala, D., & Lan, Y.-C. (2020). Team innovation in retail services: The role of ambidextrous leadership and team learning. Service Business, 14, 167–186.

[52] Ni, D., Wu, W., Zhang, Y., Chen, C., Zhang, R., & Wu, S. E. (2024). Adopting underdogs' ideas triggers fairness? When and how underachievers' voice endorsement promotes team voice. Journal of Organizational Behavior, 45, 1287–1305.

[53] Sun, Y., Yang, H., Qian, C., Jiang, Y., Luo, X., & Wu, X. (2022). Voice endorsement and employee safety voice behavior

in construction projects: The mediating role of leader-member exchange. International Journal of Environmental Research and Public Health, 19, 3374.

[54] Zinko, R., Ferris, G. R., Humphrey, S. E., Meyer, C. J., & Aime, F. (2012). Personal reputation in organizations: Two-study constructive replication and extension of antecedents and consequences. Journal of Occupational and Organizational Psychology, 85, 156–180.

[55] Hunton, J. E., Price, K. H., & Hall, T. W. (1996). A field experiment examining the effects of membership in voting majority and minority subgroups and the ameliorating effects of postdecisional voice. Journal of Applied Psychology, 81, 806–812.

[56] Johnson, H., Wu, W., Zhang, Y., & Lyu, Y. (2024). Ideas endorsed, credit claimed: Managerial credit claiming weakens the benefits of voice endorsement on future voice behavior through respect and work group identification. Human Relations, 77, 709–736.

[57] Weiss, M., Kolbe, M., Grote, G., Spahn, D. R., & Grande, B. (2018). We can do it! Inclusive leader language promotes voice behavior in multi-professional teams. The Leadership Quarterly, 29, 389–402.

[58] Azizli, N., Atkinson, B. E., Baughman, H. M., & Giammarco, E. A. (2015). Relationships between general self-efficacy, planning for the future, and life satisfaction. Personality and Individual Differences, 82, 58–60.

[59] Gyamerah, S., He, Z., Asante, D., Ampaw, E. M., & Gyamerah, E. E. D. (2022). Paternalistic leadership, employee creativity, and retention: The role of psychological empowerment. International Journal of Cross Cultural Management, 22, 83–104.

[60] Rosing, K., Frese, M., & Bausch, A. (2011). Explaining the heterogeneity of the leadership-innovation relationship: Ambidextrous leadership. The Leadership Quarterly, 22, 956–974.

[61] Vandewalle, D. (1997). Development and validation of a work domain goal orientation instrument. Educational and Psychological Measurement.

[62] Mom, T. J. M., van den Bosch, F. A. J., & Volberda, H. W. (2009). Understanding variation in managers' ambidexterity: Investigating direct and interaction effects of formal structural and personal coordination mechanisms. Organization Science, 20, 812–828.

[63] Burris, E. R. (2012). The risks and rewards of speaking up: Managerial responses to employee voice. Academy of Management Journal, 55, 851–875.

[64] Wang, C. L., & Rafiq, M. (2014). Ambidextrous organizational culture, contextual ambidexterity and new product innovation: A comparative study of UK and Chinese high-tech firms. British Journal of Management, 25, 58–76.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Impact Analysis of the Financial Shared Services Model on Mengniu Dairy's Financial Data

Hu Bo[1], Wang Ani[2]*

1. School of Management, Xi'an Polytechnic University, Xi'an, Shaanxi, 710048, China
2. Finance Department, Xianyang Caihong Optoelectronics Technology Co., Ltd., Xianyang, Shaanxi, 712023, China

Corresponding author: Wang Ani

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: The Financial Shared Services (FSS) model is a new operational framework that relies on big data platforms and internet technologies to centralize and standardize the management of accounting and reporting processes. Its core lies in achieving the centralized sharing and efficient management of financial data through the integration of unified system platforms, accounting systems, accounting standards, and operational procedures. This paper introduces the theoretical foundation of Financial Shared Service Centers (FSSCs) and uses the establishment of an FSSC by Mengniu Dairy as a case study. Based on selected data from Mengniu Dairy's financial statements from 2015 to 2024, the study examines four key financial dimensions: profitability, solvency, operational efficiency, and growth capacity. The results indicate a positive correlation between the implementation of the FSSC and changes in certain key financial metrics at Mengniu Dairy.

Keywords: Financial Shared Services Model; Financial Data; Mengniu Dairy
Published: May 10, 2026
DOI: https://doi.org/10.62177/apemr.v3i3.1404

Since the mid-to-late 20th century, with the deepening of economic globalization and the increasing internationalization of market trade, global corporate groups have faced increasingly complex challenges in their operations. Under intense market competition, many large enterprises have pursued rapid expansion through mergers and acquisitions to enhance their core competitiveness and keep pace with international trends. As international competition intensifies, subsidiaries of large conglomerates located in different regions encounter numerous challenges in financial management. To meet the management demands of this new landscape, traditional financial management models must evolve, and foundational financial operations require systematic reorganization and enhancement to achieve economies of scale. Driven by this need, the shared services model has been introduced into the financial domain by scholars and the business community, gradually gaining practical implementation. Currently, financial shared service centers are flourishing globally. Approximately 90% of Fortune Global 500 companies have established such centers, while over 70% of domestic enterprises have initiated or completed their construction. This demonstrates that building financial shared service centers not only aligns with corporate development needs but also responds to the prevailing trends of the era.

1. Theoretical and Applied Research on the Financial Shared Services Model

Financial shared services, as a novel management concept, was first proposed by scholars such as Robert Gunn (1993) and evolved from theoretical concepts to practical application in the finance domain. In the 1980s, Ford Motor Company in the

United States pioneered the adoption of the financial shared services management model [1]. By the early 21st century, the application of financial shared service centers had matured in developed economies like Europe and the United States. After a period of sustained exploration and development, research in China's financial shared services sector has gradually reached a relatively mature stage. Domestic scholars Liu Mengke and Zhu Xuanchen (2020) proposed that the fundamental purpose of enterprises establishing financial shared service centers lies in systematically simplifying and optimizing business processes. During the construction phase, digital intelligence tools are employed to reshape business scenarios, thereby strengthening the enterprise's management and control capabilities over financial activities [2]. Regarding practical application, Liu Chenyu (2024) indicates that implementing financial shared services resolves issues inherent in traditional corporate processes—such as cumbersome approvals, inconsistent standards, information opacity, and weak risk management. This facilitates a shift in financial oversight from decentralization to centralization, from manual operations to intelligent automation, and from post-event verification to end-to-end supervision, ultimately reducing operational costs [3]. Current scholarly research on the theoretical foundations of financial shared services encompasses organizational flattening theory, economies of scale theory, process reengineering theory, and standardization theory. Concurrently, literature review reveals that financial shared services exhibit the following characteristics: a customer-centric business operation model; a focus on enhancing corporate value and competitiveness; and current prevalence in large enterprises or multinational corporations.

2. Research on How the Financial Shared Services Model Enhances Corporate Financial Performance

In the current business environment, establishing a financial shared services model has become critical for corporate survival and development. Numerous scholars have demonstrated through various methodologies the positive impact of financial shared services on corporate financial performance. Hou Jing (2020) selected 81 A-share listed companies as research samples, encompassing both enterprises with and without established financial shared services, to investigate the effects of implementing such a model on corporate performance. The findings indicate that companies with financial shared services demonstrate significantly superior performance compared to those without, with a particularly pronounced effect on enhancing operational outcomes [4]. Hu Aiping (2021) argues that during the initial phase of building a financial shared service center, substantial foundational investments coupled with less immediate returns may lead to a temporary decline in financial performance. However, the long-term impact is generally positive [5]. Liu Kuinan (2021) summarized the performance-enhancing effects of the financial shared services model across four dimensions: broadening revenue streams, freeing up financial personnel, reducing operational costs, and strengthening business integration. This study highlights the significant value of financial shared service centers in optimizing corporate performance management [6].

Building upon these findings, this paper further investigates the operational mechanisms of the financial shared services model on corporate financial data. By calculating the increase or decrease in relevant financial indicators, it examines the financial changes at Mengniu Dairy before and after implementing the shared services model. Through data analysis and case studies, this paper concludes that the establishment of Mengniu Dairy's financial shared services model positively impacts the company's financial performance, providing valuable insights for enterprises considering the implementation of such models.

3. Analysis of the Impact of the Financial Shared Services Model on Mengniu Dairy's Financial Performance

This paper utilizes Mengniu Dairy's financial indicator data from 2015 to 2024 to analyze the impact of establishing a financial shared service model on financial performance across four dimensions: profitability, solvency, operational efficiency, and growth potential. The data is sourced from Mengniu Dairy's official financial statements.

3.1 Analysis of the Impact on Profitability

This paper analyzes Mengniu Dairy's profitability using three key indicators: net profit margin, return on equity, and gross profit margin. The relevant data for these indicators is presented in Table 1.

Figure 1 Profitability Indicators of Mengniu Dairy *(Unit: %)*

Project	2024	2023	2022	2021	2020	2019	2018	2017	2016	2015
Net Profit Margin	0.25	4.96	5.60	5.63	4.61	5.44	4.64	3.38	-1.51	5.14
Return on Equity	0.24	11.39	13.79	14.40	11.35	15.11	12.73	9.43	-3.50	10.85
Gross Profit Margin	39.57	37.15	35.30	36.75	37.65	37.55	37.38	35.21	32.79	31.36

Data source: Mengniu Dairy financial reports, same below.

Net profit margin refers to the percentage of net profit relative to sales revenue. This metric measures a company's ability to generate profits from sales revenue within a specific period, providing a direct reflection of the profitability level of sales revenue. Generally, a higher net profit margin indicates stronger profitability and healthier overall financial health. As shown in Table 1, Mengniu Dairy's net profit margin has trended upward since establishing its financial shared services model, peaking at 5.63% in 2021. This demonstrates that the company's digital marketing platform and financial shared services center have continuously enhanced its sales capabilities, generating greater net profits. The negative net profit margin in 2016 primarily resulted from losses incurred by its holding company, Yashili International Group Co., Ltd., rather than losses from its core operations. Additionally, the sharp decline to 0.25% in 2024 stemmed from a significant goodwill impairment charge of RMB 4.33 billion. Although the net profit margin fluctuated in other years, overall, the net profit margin after implementing the financial shared services model was significantly higher than before its establishment. This outcome demonstrates that building the financial shared services model has played a positive role in enhancing the company's net profit margin. Return on equity (ROE) is the ratio of net profit to average shareholders' equity, reflecting the return level of shareholders' equity to some extent and measuring the efficiency of capital utilization. A higher value indicates greater returns on investment. As shown in Table 1, Mengniu Dairy's return on equity in 2020 experienced a significant decline compared to 2019, primarily due to the impact of the COVID-19 pandemic. Nevertheless, this metric remained above the level prior to the establishment of the financial shared service model. This demonstrates that the model has helped maintain Mengniu Dairy's return on equity within a relatively high range, enabling the company to sustain profitability and maintain a stable and positive overall development trajectory. Gross profit margin is a core financial indicator measuring a company's profitability in its primary business operations, reflecting the inherent profit potential of its products or services. As shown in Table 1, Mengniu Dairy maintained a gross profit margin above 30% from 2015 to 2023. This demonstrates that the implementation of the financial shared services model significantly enhanced the company's brand recognition and expanded its market share.

3.2 Analysis of the Impact on Debt Repaying Capacity

This paper analyzes Mengniu Dairy's debt-paying capacity using two indicators: the current ratio and the debt-to-equity ratio. The relevant indicators are shown in Table 2.

Figure 2 Mengniu Dairy's Debt Repayment Capacity Indicators

Project	2024	2023	2022	2021	2020	2019	2018	2017	2016	2015
Current Ratio	0.97	1.05	1.11	1.07	1.21	1.18	1.18	1.15	1.30	1.40
Asset-Liability Ratio (%)	54.82	56.04	57.52	57.19	53.53	57.74	54.16	53.38	48.08	47.46

The current ratio is the ratio of a company's current assets to its current liabilities. A higher value indicates stronger liquidity and superior short-term debt repayment capacity; conversely, a lower value signifies weaker short-term debt repayment ability. As shown in Table 2, since initiating its financial shared services model in 2015, Mengniu Dairy's current ratio has generally trended downward. The core reason lies in the substantial capital investment required to implement this model, which caused temporary fluctuations in cash flow. After 2018, Mengniu's Financial Shared Service Center achieved simultaneous financial accounting and fund settlement across all business modules and processes. This significantly enhanced the precision of fund management, gradually stabilizing and optimizing the current ratio from its initial downward trend during the construction phase. The debt-to-asset ratio represents the percentage of total liabilities relative to total assets at the end of a period, reflecting a company's ability to leverage creditor-provided funds for operations. Table 2 further

shows that Mengniu Dairy's debt-to-asset ratio rose from 47.46% in 2015 to 57.74% in 2019, exceeding the 50% threshold. This increase primarily resulted from Mengniu Dairy's ongoing advancement of its financial shared services model and increased investment in digital marketing, which elevated expenses related to advertising and promotion. The subsequent stabilization trend stems mainly from Mengniu establishing a big data foundation through its financial shared services center. This foundation assists managers in identifying operational improvement opportunities, further enhancing asset utilization efficiency and promoting a healthier debt-to-asset ratio.

3.3 Analysis of Impact on Operational Capability

This paper analyzes Mengniu Dairy's operational efficiency using two key indicators: total asset turnover and inventory turnover. The relevant metrics are presented in Table 3.

Figure 3 Mengniu Dairy Operating Capability Metrics *(Unit: Times)*

Project	2024	2023	2022	2021	2020	2019	2018	2017	2016	2015
Total Asset Turnover Ratio	0.99	0.85	0.86	0.99	0.96	1.09	1.11	1.12	1.08	1.00
Inventory Turnover Ratio	9.72	8.75	8.23	9.29	8.94	10.53	11.09	11.42	9.44	7.75

Total asset turnover is a core financial indicator measuring a company's asset operational efficiency. It reflects the ability of all assets to generate operating revenue within a specific period, demonstrating the speed and utilization efficiency of assets from input to output. As shown in Table 3, Mengniu Dairy's total asset turnover rate exhibited a sustained upward trend from 2015 to 2017, indicating positive performance. This trend clearly demonstrates that the financial shared services model effectively enhances the company's asset utilization efficiency, confirming its strong operational quality. However, impacted by the COVID-19 pandemic, Mengniu Dairy's total asset turnover ratio has generally trended downward since 2018. This shift may stem from the company's sustained increase in capital investment, driving continuous expansion of total assets. Yet, as these investments failed to generate timely returns, asset turnover efficiency ultimately declined. Inventory turnover rate, defined as the ratio of operating costs to average inventory balance over a specific period, reflects the speed at which inventory is turned over. Data in Table 3 further reveals that Mengniu Dairy's inventory turnover rate showed a sustained upward trend from 2015 to 2017. This indicator change clearly demonstrates that the financial shared services model has effectively enhanced the company's inventory management capabilities, significantly accelerating inventory turnover speed. Affected by the COVID-19 pandemic, Mengniu Dairy's inventory turnover rate experienced a sharp decline in 2020 compared to 2019 and remained at a relatively low level in subsequent years. Subsequently, due to the establishment of a slow-moving inventory disposal team led by the shared service center, over 100 million yuan worth of unsold products were cleared cumulatively from 2021 to 2023, releasing substantial funds and enabling an improvement in the inventory turnover rate for the 2024 fiscal year.

3.4 Analysis of the Impact on Growth Potential

Figure 4 Mengniu Dairy Growth Capability Indicators *(Unit: %)*

Project	2024	2023	2022	2021	2020	2019	2018	2017	2016	2015
Operating Revenue Growth Rate	-10.09	6.51	5.05	15.92	-3.79	14.57	14.66	11.86	9.69	-2.04
Total Assets Growth Rate	-7.75	-2.21	20.00	22.40	2.05	18.18	14.31	18.36	-3.02	7.06

The calculation logic for revenue growth rate is the ratio of the current year's revenue increase to the previous year's total revenue. This metric uses the company's historical revenue scale as a benchmark, providing a direct measure of the current period's revenue growth rate and development trajectory. Guided by the financial shared services model, Mengniu Dairy's revenue growth rate rose from -2.04% in 2015 to 9.69% in 2016, successfully shifting from negative to positive territory. Affected by the COVID-19 pandemic, Mengniu Dairy's revenue growth rate plummeted in 2020. Subsequently, the company implemented countermeasures, advancing its digital transformation strategy alongside the shared services model, which restored the revenue growth rate to an upward trajectory. Overall, the financial shared services model has driven revenue

growth for Mengniu Dairy, providing enhanced development opportunities. The total asset growth rate serves as a core financial indicator reflecting the pace of corporate asset expansion and development potential. It measures the increase in total assets over a specific period, directly illustrating a company's expansion capacity and resource accumulation speed. As shown in Table 4, the implementation of the financial shared services model from 2015 to 2023 led to a total asset growth rate that initially slowed before accelerating. The initial slowdown was due to system integration investments. After 2017, as the platform matured, growth began to accelerate. The implementation of the financial shared services model enhanced the company's ability to centrally allocate funds, reduced idle capital, improved capital turnover rates, and indirectly increased the efficiency of total asset utilization. Overall, the financial shared services model transformed Mengniu's total asset growth from extensive to steady, high-quality expansion, while shifting its operating revenue growth from volatility to sustainability. This provides a model case for large dairy groups to achieve strategic upgrades through financial management innovation.

4. Conclusions and Limitations

Driven by the wave of the digital economy, modern information technology has entered an era of rapid iteration and deep application. In China, the theoretical research and commercial practice of financial shared services have been cultivated for over three decades. Leveraging core advantages of digitalization, centralization, and standardization, the establishment and optimization of financial shared services platforms have increasingly become a crucial lever for enterprises to advance the transformation and upgrading of financial management and achieve overall operational digitalization and intelligence. This approach has garnered widespread market favor and proactive implementation. For fast-moving consumer goods like dairy products, real-time information transmission is particularly critical. The financial shared services model can effectively reduce the risks of product expiration and slow sales by breaking down data silos and enabling efficient information flow. Therefore, this paper analyzes Mengniu Dairy's profitability, solvency, operational efficiency, and growth potential across four dimensions, revealing that establishing a financial shared services model can promote its development to a certain extent. Mengniu Dairy should fully leverage surrounding resources to better utilize its advantages throughout production processes and the industrial chain. It should also adjust the implementation form of the financial shared services model based on its development status at different stages to achieve high-quality development. However, this study has limitations. Research on the correlation between financial shared services and corporate financial data requires further refinement. Additionally, data collection faced challenges due to Mengniu Group's status as a Hong Kong-listed company, whose disclosure rules differ from mainland Chinese listed companies, making it difficult to ensure the accuracy of some data.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Gunn, R. W., Carberry, D. P., Frigo, R., & Behrens, S. (1993). Shared services: Major companies are reengineering their accounting functions. Management Accounting, 75(5), 22–29.

[2] Liu, M., & Zhu, X. (2020). Practice and service model innovation of financial shared service centers in real estate groups under informationization: A case study of Group B. Construction Economics, 41(S2), 41–45.

[3] Liu, C. (2024). Implementation plan for process reengineering under China Railway Construction Corporation's shared services model. Finance and Accounting, (11), 64–65.

[4] Hou, J. (2020). Study on the impact of shared financial services in listed companies on corporate performance (Master's thesis). Xi'an University of Science and Technology.

[5] Hu, A., Zhang, C., & Zhou, S. (2021). Research on the impact of financial shared services on corporate financial performance: A case study of Huawei Company. Accounting Friend, (19), 14–19.

[6] Liu, K. (2021). Performance management of financial shared service centers based on cloud accounting in the big data era. Accounting Learning, (07), 99–100.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Optimization of University Financial Management Models Empowered by Artificial Intelligence

Fei Chen*

Finance Office, Nantong University, Nantong, Jiangsu, 226019, China

Corresponding author: Fei Chen

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Against the backdrop of the modernization of university governance, the digital transformation of education, and the continued development of smart campuses, financial management in higher education institutions is shifting from a traditional accounting- and support-oriented model toward one that is more intelligent, collaborative, and governance-oriented. However, in practice, university financial management still faces a number of challenges, including insufficient system coordination, cumbersome operational processes, limited data utilization capacity, and increasing pressure in risk prevention and control, all of which constrain further improvements in financial management effectiveness. Based on an examination of the current development status and major challenges of university financial management, this paper analyzes the mechanisms and practical significance of AI empowerment in this field, arguing that artificial intelligence can play an important role in improving operational efficiency, strengthening decision-support capacity, and enhancing risk prevention and control. On this basis, the paper further identifies the objectives and key directions for optimizing university financial management models under AI empowerment, emphasizing the transition of financial management from traditional transactional processing to modern governance support. It also proposes implementation paths from the perspectives of data governance, platform development, and process reengineering, while outlining supporting mechanisms in terms of institutional regulation, interdepartmental coordination, and evaluation and feedback. The study concludes that AI-empowered financial management should be integrated into the overall digital governance framework of universities, and that sustained improvements in financial governance capacity and management performance can be achieved only through the coordinated advancement of technological application, process optimization, and institutional innovation.

Keywords: University Finance; Artificial Intelligence; Optimization of Management Models; Digital Governance

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1403

1. Current Development Status and Major Challenges of University Financial Management

In recent years, against the background of reforms in budget management, strengthened performance management, and accelerated smart campus development, university financial management has gradually shifted from the traditional accounting-oriented model toward greater standardization, informatization, and refinement, with its overall management level steadily improving. Relying on financial information systems, some universities have incorporated such functions as budget management, accounting, reimbursement approval, and fee management into online operations. Application scenarios such

as e-invoices, online reimbursement, and financial information inquiry have continued to expand, significantly improving the convenience and service efficiency of financial management [1]. Nevertheless, university financial management still faces many practical problems and challenges.

To begin with, the problems of fragmented financial data and insufficient system coordination remain prominent. Although many universities have established multiple business systems, the data standards, interface rules, and operating logic among these systems have not yet been fully unified. Data sharing between the financial system and systems for personnel, research, assets, procurement, and academic affairs remains limited, and information silos still exist. This not only affects data integration and utilization, but also constrains the transformation of financial management toward comprehensive analysis and coordinated governance.

In addition, financial management processes in universities remain relatively cumbersome. Influenced by traditional management models, some universities still face such problems as too many approval levels, repeated reviews, and strong manual intervention in reimbursement approval, fund auditing, budget adjustment, and project settlement. As a result, financial operations remain inefficient, and the service experience for faculty and students still needs improvement. Especially in areas such as research fund management and the supervision of special-purpose funds, financial personnel often have to devote substantial effort to manual verification and procedural review, which easily leads to a heavy burden of transactional work and insufficient performance of management functions.

Moreover, the data analysis and decision-support capacity of university financial management remains inadequate. Although many universities have accumulated a certain scale of financial data, applications in data mining, trend analysis, risk identification, and performance evaluation remain limited. Financial data are still largely confined to recording, statistics, and aggregation, and have not yet been fully transformed into effective support for resource allocation optimization, cost structure analysis, and management decision-making. To some extent, this constrains the expansion of financial management from result reporting to process management, and from ex post summarization to forward-looking prediction.

Furthermore, university financial management is also under new pressure in internal control and risk early warning. With the expansion of university operations, increases in funding volume, and increasingly diversified sources of funds, financial risks have become more complex, and the requirements for risk identification and prevention have become significantly higher. Traditional management methods that rely on manual review and experience-based judgment often suffer from delayed identification, insufficient early warning, and untimely response when confronted with high-frequency, complex, and multi-scenario financial activities, and are no longer able to meet the practical needs of refined financial governance.

Overall, university financial management has already established a certain developmental foundation, and positive progress has been made in information infrastructure, institutional regulation, and management philosophy. However, it still faces major challenges in system coordination, process optimization, data application, and risk prevention and control. This indicates that although university financial management has entered the stage of digital transformation, it still falls short of a more efficient, precise, and intelligent modern financial management model. Therefore, actively introducing new technologies such as artificial intelligence and promoting the optimization and upgrading of university financial management models has become an important direction for improving university financial governance capacity and management performance.

2. Mechanisms and Practical Significance of AI Empowerment in University Financial Management

As an important driving force of the new round of scientific and technological revolution and industrial transformation, artificial intelligence is profoundly reshaping organizational operation modes and resource allocation patterns. For university financial management, AI is not only a technical tool for improving work efficiency, but also an important force for promoting management model optimization and governance capacity enhancement. With the continuous development of big data, machine learning, natural language processing, and intelligent recognition technologies, application scenarios of AI in university financial management are becoming increasingly rich, and its influence is gradually extending from basic business assistance to deeper areas such as process optimization, risk prevention and control, and decision support [2].

In terms of its mechanism of action, the empowering effect of AI on university financial management is first reflected in

the improvement of financial operational efficiency. University financial management involves a large amount of repetitive and procedural work, such as invoice review, reimbursement approval, data entry, and information verification. Traditional approaches rely heavily on manual operations, which not only prolong processing cycles but also tend to generate errors. With the help of intelligent recognition, process automation, and rule-matching technologies, AI can rapidly handle part of these basic affairs-like tasks, reduce the manual workload, and improve the speed and accuracy of financial business processing.

AI empowerment is also reflected in the enhancement of decision-support capacity. University financial systems accumulate large amounts of data on budget execution, fund utilization, cost distribution, and asset allocation. However, such data can become true governance resources only after effective integration and in-depth analysis. AI can identify patterns, detect deviations, and predict trends from massive volumes of data, thus providing a more scientific basis for budget preparation, resource allocation, performance evaluation, and fund utilization optimization, and promoting a shift in financial management from experience-based judgment to data-driven decision-making.

At the same time, AI empowerment is also manifested in the strengthening of risk prevention and control. University financial activities are diverse in type and frequent in capital flows, while traditional supervision methods relying on manual review often suffer from limitations such as delayed identification and insufficient coverage. By relying on model analysis and intelligent comparison, AI can timely identify and dynamically warn against risk points such as abnormal invoices, irregular expenditures, and budget deviations, thereby improving the precision and forward-looking nature of internal control and enhancing the proactiveness of financial supervision.

From the perspective of practical significance, AI empowerment in university financial management carries evident value. On the one hand, it helps improve the efficiency and service quality of financial management, alleviates the transactional workload of financial personnel, and enables them to devote more energy to management analysis and value creation. On the other hand, it helps improve the utilization efficiency of financial data, enhances the scientific basis of budget management, performance evaluation, and resource allocation, and provides stronger data support for university governance. Meanwhile, AI also helps strengthen internal control and risk prevention, promotes the shift of financial supervision from ex post review to whole-process monitoring, and further facilitates the transformation of financial management functions from traditional accounting support to management service and governance support.

Overall, what AI brings to university financial management is not merely efficiency improvement at the technical level, but also a systematic transformation in management philosophy, operational methods, and governance capacity. Through automation, intelligence, and data-driven approaches, AI provides important support for the optimization of university financial management models and lays the foundation for clarifying optimization objectives and identifying key directions in subsequent sections.

3. Objectives and Key Directions for Optimizing University Financial Management Models Under AI Empowerment

On the basis that the mechanisms and practical significance of AI empowerment in university financial management have become increasingly clear, further clarifying the objectives and key directions of financial management model optimization has become an important link in promoting the transformation and upgrading of financial management. Optimization under AI empowerment should not merely remain at the level of improving business processing efficiency and reducing transactional burdens, nor should it be confined to local process improvements and single-item technological applications. Rather, it should focus on the restructuring of financial functions, the transformation of management methods, and the enhancement of governance capacity, thereby promoting the transformation of university financial management from a traditional accounting-support model to an intelligent collaborative and governance-support model [3].

In terms of objective positioning, the optimization of university financial management models under AI empowerment should place greater emphasis on governance orientation. Traditional university financial management has focused more on accounting, fund review, and rule implementation, emphasizing compliance and accuracy. Under the background of digital and intelligent development, however, the functional boundaries of financial management are continuously expanding. Financial departments should not only undertake the responsibility of fund support, but also play a more active role in

budget performance management, resource allocation optimization, risk identification and early warning, and management decision support. In other words, the core of optimizing university financial management models lies not merely in improving processing speed and technical capability, but in using AI to promote a shift in financial management from "bookkeeping and reimbursement handling" to "analysis and judgment," and from "ex post supervision" to "whole-process governance."

On this basis, the optimization of university financial management models should focus on such directions as process coordination, data-driven governance, service enhancement, and risk control. In terms of process coordination, financial operations should be transformed from fragmented handling to intelligent collaboration. University financial management involves multiple links such as budget preparation, reimbursement approval, fund management, asset allocation, and performance evaluation. Under the traditional model, these links often suffer from fragmented processes, disconnected procedures, and insufficient coordination. AI-empowered financial management should therefore place greater emphasis on process reshaping and business collaboration, promoting financial matters from isolated processing points to whole-process linkage, and continuously improving the integrity and operational efficiency of management.

In terms of data application, financial management should be transformed from experience-based judgment to data-driven governance. As university operations become increasingly complex, financial management can no longer rely solely on experience-based judgment and static statistics, but should instead be based on fuller data integration and intelligent analysis. Through AI empowerment, optimization of university financial management models should pay greater attention to the correlation analysis between financial data and business data related to personnel, research, assets, and procurement, thereby continuously improving the scientific basis of budget arrangement, cost control, performance evaluation, and resource allocation, and turning financial data into an important resource supporting university governance.

In terms of service functions, financial management should shift from passive response to precise services. Traditional financial management has emphasized procedural review and institutional constraints, and service provision has been relatively single and unable to fully meet the diversified and convenient needs of faculty, students, and administrative departments. With AI support, the optimization of university financial management models should place greater emphasis on service awareness and service capability, continuously improving the timeliness, convenience, and relevance of financial services through intelligent Q&A, information reminders, process guidance, and personalized feedback.

In terms of risk governance, financial supervision should be transformed from ex post control to whole-process risk prevention and control. Faced with the new situation of expanding fund scales, growing business scenarios, and rising regulatory requirements, traditional supervision methods that rely solely on ex post review and experience-based judgment can no longer meet practical needs. AI-empowered financial management model optimization should integrate risk identification, early-warning monitoring, and dynamic control throughout budget execution, fund utilization, invoice review, and fund supervision, thereby continuously enhancing the forward-looking, precise, and continuous nature of internal control.

Overall, the objective of optimizing university financial management models under AI empowerment is to promote the transformation of financial management from traditional transactional processing to modern governance support, while the key directions are concentrated in process coordination, data-driven governance, precise services, and whole-process risk prevention and control. These directions not only respond to the practical problems currently faced by university financial management, but also provide clear guidance for the subsequent design of implementation paths and supporting mechanisms.

4. Implementation Paths and Safeguard Mechanisms for Optimizing University Financial Management Models Under AI Empowerment

On the basis that the objectives and key directions have become basically clear, the key to transforming the conceptual design of AI-empowered university financial management model optimization into actual results lies in forming clear and feasible implementation paths, together with corresponding safeguard mechanisms to ensure stable operation and sustained deepening. The optimization of university financial management models cannot remain at the level of local technological embedding and single-function improvement. Rather, it should be advanced in a coordinated way around data governance, platform development, process reengineering, and talent support, while being safeguarded through institutional regulation, coordinated linkage, and evaluation and feedback mechanisms.

From the perspective of implementation paths, data governance is the fundamental link in AI-empowered university financial management. The effective functioning of AI depends to a large extent on data quality and data connectivity. Therefore, universities should strengthen the standardization of financial data, promote data sharing and business coordination between financial systems and systems for personnel, research, assets, procurement, and academic affairs, gradually break down information silos, and improve the capacity for data integration and utilization. At the same time, they should strengthen the classified management, dynamic updating, and standardized use of financial data so as to provide authentic, complete, and usable data support for budget preparation, performance evaluation, risk early warning, and management decision-making [4].

On the premise of consolidating the data foundation, platform development serves as an important support for the practical application of AI. Universities should build on existing financial informatization conditions, promote interconnection among business platforms for budget management, accounting, reimbursement approval, fee management, asset management, and performance evaluation, and gradually establish an intelligent platform system covering the whole process of financial management. By integrating fragmented systems, optimizing functional modules, and expanding application scenarios, universities can promote the in-depth application of AI in invoice recognition, intelligent review, anomaly warning, data analysis, and service consultation, thereby continuously improving the operational efficiency and collaborative capacity of financial management.

Process reengineering, in turn, is the key lever for achieving model optimization. AI empowerment in university financial management does not simply mean replacing manual work with technology; rather, it means using technological means to reshape business processes. In response to problems such as lengthy reimbursement approval chains, repetitive review procedures, and poor cross-departmental coordination, universities should systematically sort out financial business processes, reduce unnecessary intermediate links, optimize approval logic and processing routes, and promote financial matters from fragmented handling to whole-process coordination, so that technological application can truly be transformed into improvements in management performance.

Meanwhile, the in-depth application of AI in university financial management also relies on corresponding talent support. Universities should therefore continuously improve the data analysis, information technology application, and collaborative management capabilities of financial personnel through special training, on-the-job practice, and interdisciplinary cooperation, promoting their transformation from traditional business executors into participants in intelligent finance and supporters of governance, thereby providing sustained momentum for AI-empowered financial management.

From the perspective of safeguard mechanisms, the first requirement is to improve the institutional regulatory system. The application of AI in university financial management must be grounded in sound institutional development. Universities should establish and improve relevant institutional systems concerning data usage permissions, business operation standards, risk control requirements, and boundaries of responsibility, so as to ensure that AI applications are advanced on a compliant, secure, and controllable track [5].

The second requirement is to improve cross-departmental coordination mechanisms. AI empowerment in university financial management cannot be accomplished by the financial department alone; rather, it requires the joint participation and coordinated advancement of departments such as finance, information technology, research, personnel, assets, and auditing. Universities should establish linkage mechanisms with clear responsibilities, reasonable division of labor, and smooth communication, thereby forming a work pattern characterized by vertical connection and horizontal coordination and providing strong organizational support for the optimization of financial management models.

The third requirement is to establish evaluation and feedback mechanisms. Whether AI applications can truly be transformed into management effectiveness must ultimately be tested through practice. Universities should establish a scientific evaluation system around such aspects as application effectiveness, service quality, risk control, and governance outcomes, and should promptly adjust technological application and management processes according to evaluation results, so as to continuously deepen AI empowerment in university financial management.

Overall, the implementation paths for optimizing university financial management models under AI empowerment focus on solving the problem of “how to advance implementation,” while the safeguard mechanisms focus on solving the problem

of "how to ensure stable operation." Only by organically integrating implementation paths with safeguard mechanisms can universities truly promote the in-depth application of AI in financial management and continuously improve their financial governance capacity and management performance.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Xu, C. (2025). Analysis on the optimization of university financial sharing in the era of digital intelligence. Digital Communication World, (06), 244–246.

[2] Xu, N., & Zeng, F. (2024). Research on the optimization and application of artificial intelligence algorithms in university financial management. China Informatization, (11), 93–95.

[3] Sun, S. (2025). Research on the optimization of university financial control models in the age of artificial intelligence. Landbridge Horizon, (08), 57–59.

[4] Ma, L. (2025). Discussion on optimization strategies for university financial management from the perspective of big data. Guangdong Economy, (20), 74–76.

[5] Hong, B. (2026). Research on the optimization path of university financial management under the background of big data. Chinese Agricultural Accounting, 36(03), 59–61.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Enhancement of Technological Transformation Efficiency under New Quality Production Forces

Wendi Cui*

School of Management, Xi'an Polytechnic University, Shaanxi, Xi'an, 710048, China

Corresponding author: Wendi Cui, cuiwendi293@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: The cultivation and development of new quality productive forces rely on the deep integration of scientific and technological innovation and industrial innovation. Currently, the scientific and technological achievement transformation system in our country is facing deep obstructions such as fragmented resource elements, coordination barriers, and insufficient compatibility of incentive systems. Innovation is disconnected from industry, unable to achieve efficient interaction between science and the market domain, resulting in the inability to transform scientific and technological achievements into real productive forces. Based on the background of new quality productive forces, this paper analyzes the current status of scientific and technological achievement transformation, addresses the identified problems, and proposes corresponding measures to provide strong support for improving the efficiency of the scientific and technological achievement transformation system and promoting and cultivating new quality productive forces.

Keywords: New Quality Productive Forces; Scientific and Technological Achievement Transformation; Efficiency Improvement
Published: May 10, 2026
DOI: https://doi.org/10.62177/apemr.v3i3.1411

Introduction

In the current era of rapid global technological advancement, scientific and technological innovation has become the core driving force for promoting high-quality development of the economy and society. Unlike traditional productivity and general productivity, the process of new quality productivity's incubation encounters various key core technology breakthroughs[1] and "knockout" type innovation difficulties and pain points[2]. Especially for strategic emerging industries and future industries that serve as the main frontlines, the required original innovation often involves large investment, long cycles, and high risks[2]. On one side is scientific and technological innovation, and on the other is industrial development. There exists an invisible "gap" between the two, and this "transformation gap" reflects not only the efficiency issue of technology transfer, but also the structural predicament of the innovation system. The traditional thinking of technological transformation regards scientific and technological innovation as the source and industrial innovation as the result, attempting to achieve the leap from the laboratory to the production line through linear recursion. However, this simplified model has significantly reduced its explanatory power and guiding force in today's era of accelerated technological iteration, interdisciplinary integration, and complex and changeable markets.

New quality productivity itself is a productivity form where innovation plays a leading role. It requires breaking the

institutional barriers of "two skins" between science and economy, and establishing a new type of relationship where the innovation chain and the industrial chain are coordinated. This article aims to transcend the traditional cognitive framework of "linear transformation", and re-examine the efficiency improvement of the technology transformation system from the perspective of system coupling. Therefore, this article first analyzes the problems currently faced by technology transformation in the Chinese context, then explores the deep reasons for these problems, and subsequently proposes systematic improvement measures oriented towards the development of new quality productivity, providing theoretical references and lessons for enhancing the efficiency of technology transformation and promoting the development of new quality productivity.

1. Structural Issues of the Technology Transformation System

1.1 Structural Break in the Innova2tion Chain

From the laboratory prototype to industrial application, scientific and technological achievements need to go through multiple stages such as basic research, application development, pilot-scale optimization, and large-scale production. Currently, in China's innovation chain, the basic research stage is supported by scientific funds, and the industrial application stage is covered by industrial policies. However, the intermediate area connecting these two stages has obvious institutional vacuum and resource shortage. This "death valley" causes a large number of scientific and technological achievements with potential new productive forces to die prematurely when crossing the gap from the laboratory to the market. Statistics show that although China has the highest number of patent applications globally, the patent conversion rate has remained below 10% for a long time, and the proportion of truly formed products is less than 5%. In the end, it is not that the scientific and technological achievements themselves lack value, but that there is a lack of engineering capabilities to transform laboratory principles into producible products. The absence of intermediate organizations such as pilot-scale bases and concept verification centers has led to a large number of innovative achievements stopping at papers and patent certificates, unable to enter the industrial cycle and be transformed into actual productive forces.

1.2 Structural Imbalance of the Mainstream Relationship

The transformation of scientific and technological achievements involves various entities such as universities, research institutions, enterprises, government agencies, and intermediary organizations. These entities should form a collaborative and interactive innovation network. However, in reality, there is a significant structural imbalance among the entities. From the perspective of research objectives, the research evaluation systems of universities and research institutions have long focused on papers and awards as core indicators, leading to research topics deviating from the actual market demands. Researchers pursue recognition within the academic community rather than market-based matching, and there is a mismatch between the output of research results and industrial demands. From the perspective of capability distribution, universities and research institutions possess abundant innovative source resources but lack industrialization capabilities. Small and medium-sized enterprises have market acumen but lack the ability to identify and digest technologies. Leading enterprises, although having the transformation capabilities, tend to conduct internal research rather than external absorption. The mismatch of capabilities among the entities makes various innovation resources exist in different entities, making it difficult to cross organizational boundaries. From the perspective of power structure, administrative forces still dominate the resource allocation process, and the decisive role of the market mechanism has not been fully exerted. Government-led scientific and technological projects, park construction, and platform establishment often have a lag and deviation from market demands, resulting in redundant resource allocation and ineffective accumulation.

1.3 Lack of organizational structuring of the service system

An efficient transformation system requires specialized and market-oriented service organizations as its support. However, the current technology intermediary service system in China has obvious organizational structuring deficiencies. From the organizational form perspective, technology transfer institutions generally have the problems of "small, scattered, and weak". Most institutions remain at the shallow level of connecting and facilitating, lacking the ability to provide deep-level services such as technical assessment, legal and financial affairs, pilot production and maturation, and business planning. From the talent structure perspective, there is a serious shortage of comprehensive technical managers who understand technology,

law, market, and finance. The number of certified technical managers across the country is less than ten thousand, and the gap compared to market demand is huge. From the platform layout perspective, various technology markets, innovation stations, and technology transfer centers in different regions are redundantly constructed, but they lack interconnection and fail to form a unified national technology factor market network. Enterprises do not know where to find the technology when they have demands, and universities do not know where to promote their achievements. The lack of information symmetry severely restricts the optimization allocation of innovation resources. The organizational structuring deficiencies of the service system have led to the fact that scientific and technological achievements lack professional intermediaries during the process of moving from the laboratory to the market, resulting in high transaction costs and difficult improvement of transformation efficiency.

2. Analysis of the Deep Causes of the Problem

2.1 Conflict of Multiple Goals

The technology transfer system has multiple goals simultaneously, including scientific exploration and economic services, state-owned capital security and market vitality, public interests and private property rights. These pairs of goals are not naturally coupled but rather have profound internal tensions. The realization path often requires seeking dynamic balance under multiple constraints. The bureaucratic logic pursues procedural compliance, risk avoidance, and clear responsibilities, reflecting strict approval procedures, rigid assessment indicators, and clear accountability mechanisms. The market logic pursues efficiency priority, compatibility of incentives, and survival of the fittest, manifested as autonomous decision-making, risk assumption, and shared benefits. In the context of technology transfer, these two logics often lead to institutional design dilemmas and distorted operational mechanisms. Taking the transfer of scientific and technological achievements by positions as an example, the state-owned assets management department requires value assessment of the achievements to prevent loss, but the value of scientific and technological achievements has high uncertainty, and the assessment results are difficult to reflect the true market expectations. Researchers hope to make quick decisions to seize the market window, but the management must fulfill the approval procedures for the disposal of state-owned assets. As a result, the procedures are compliant, but efficiency is sacrificed; form is transformed, but substance is stagnant. This conflict and compromise of multiple goals is the deep institutional root of the low efficiency of the technology transfer system.

2.2 Delay in the Development of Factor Markets

The transformation of scientific and technological achievements is essentially a process of reconfiguration of innovative factors, which requires a well-developed market mechanism as a foundation. However, the development of the innovation factor market is currently significantly lagging behind the reform processes of the product market and the capital market. In the technology market, the non-standardization of technological achievements, the uncertainty of value, and information asymmetry all pose high transaction costs for technology transactions. However, the existing technology market still has significant deficiencies in the construction of mechanisms such as price discovery, credit guarantee, and dispute resolution, failing to effectively reduce transaction costs. In the talent market, the channels for researchers to flow to enterprises or the frontlines of industries are not sufficiently smooth. The institutional designs of universities and research institutions, such as staffing management, professional title evaluation, and social security, result in high costs and significant risks for researchers to leave the system. The structural defects of the financial market are also obvious, lacking specialized financial institutions that can effectively identify and price technological risks. The lag in the development of the innovation factor market has led to the lack of effective market signals for technology transformation and support from transaction mechanisms, making it difficult to achieve optimal allocation of factors.

2.3 Structural Constraints on Cognitive Abilities

Technology transfer is a highly complex and systematic activity that requires participants to possess cross-disciplinary knowledge and capabilities. However, in reality, the cognitive structures and ability endowments of various entities exhibit significant structural constraints. From the supply side, researchers have long specialized in specific academic fields and have a deep understanding of the technology itself, but they lack systematic knowledge of market demands, business models, and intellectual property operation. They tend to overestimate the maturity of the technology and underestimate the difficulty of

industrialization, and they lack an objective assessment of the commercial prospects of the results. From the demand side, business managers are familiar with production and operation, but they lack the ability to identify and evaluate cutting-edge technologies, and they have difficulty detecting and seizing potential technological opportunities at an early stage. From the service side, the practitioners of intermediary institutions mostly lack a background in science and engineering, have limited understanding of the technology itself, and are unable to provide in-depth professional services. This knowledge gap hinders effective communication between the supply and demand sides, makes it difficult to accurately assess the value of the technology, and makes it hard to promptly capture cooperation opportunities. The structural constraints on cognition and ability are the micro-level foundation for the inefficient operation of the technology transfer system.

3. Policies for Enhancing the Efficiency of the Technology Transfer System

3.1 Deepen Institutional Innovation to Resolve the Incompatibility of Incentives

To address the issues of insufficient property rights incentives and the absence of a risk-sharing mechanism, greater efforts are needed to promote institutional innovation and establish an incentive-compatible mechanism that is compatible with the cultivation of new productive forces. At the property rights system level, comprehensive reforms of the ownership or long-term usage rights of scientific and technological achievements should be implemented. Based on the summary of pilot experiences, the coverage of the empowerment reform should be expanded, and the ownership of scientific and technological achievements or a usage right of not less than ten years should be granted to researchers. The list of exemption from responsibility for due diligence should be improved. As long as the transformation process is compliant and the pricing procedure is fair, the subsequent fluctuations in market value should not be held accountable to the managers and researchers. At the income distribution level, a market-based income-sharing mechanism should be established. The proportion of rewards for scientific researchers from the income of scientific and technological achievements transferred by universities and research institutions should be no less than 70%, and researchers should be allowed to participate in the distribution of enterprise income through equity investment based on the value of the technology. For technology transfer institutions and technology managers, a salary incentive mechanism linked to the transformation effectiveness should be established to attract and retain high-level professional talents.

3.2 Reconstruct the Innovative Organizational Model and Fill the “Death Valley”

In order to solve the problems of innovation chain breakage and the lack of intermediate links, it is necessary to start from organizational innovation and build a support system covering the whole cycle of achievement transformation. On the front end of the results, support the construction of a number of proof-of-concept centers focusing on specific areas. The task of these centers is to carry out preliminary verification of technical feasibility and commercial potential of laboratory results, and help promising results cross the first threshold from technology to product. The proof-of-concept center should be led by universities and institutes, jointly built by leading enterprises and investment institutions, operated in a market-oriented way, and the government should give start-up funds and operating subsidies. In the middle end of the achievements, the layout and construction of a number of open and shared market-oriented pilot test platforms. Relying on the national high-tech zone, economic development zone and other industrial clusters, the pilot test platform with the functions of engineering technology development, sample trial production and process verification will be built. The platform should adopt the mode of “enterprise main body, market operation and government guidance” to provide low-cost engineering services for universities, institutes and small and medium-sized enterprises. On the back end of achievements, we will support leading enterprises to lead the formation of innovation consortias, carry out collaborative research around key links of the industrial chain, and accelerate the large-scale application and iterative upgrading of technological achievements.

3.3 Cultivate a Specialized Service Ecosystem and Reduce Transaction Costs

To address the fragmentation of the service system and the lack of professional capabilities, it is necessary to focus on the direction of specialization and marketization, and systematically cultivate a technology transfer service ecosystem. At the level of talent development, efforts should be made to accelerate the cultivation of a high-quality team of technology managers. Incorporate technology managers into the national occupational classification dictionary and establish a professional title evaluation channel. Support universities in establishing master’s degree programs in technology transfer to

cultivate interdisciplinary talents who are proficient in both technology and management. Encourage industry associations to organize specialized training and qualification recognition, and establish a working mechanism where technology managers are involved throughout the process of technology transfer. At the institutional capacity building level, focus on enhancing the professionalism of technology transfer institutions. Support universities and research institutions in establishing specialized technology transfer institutions with independent legal person status, operate according to market principles, and establish flexible salary incentive systems. Encourage social forces to establish professional technology transfer service companies, forming a market-oriented and diversified service supply pattern. At the platform construction level, utilize technologies such as big data and artificial intelligence to create a national integrated technology element market network. Promote the interconnection of science and technology markets and technology trading platforms across regions, establish a national technology achievement information database, unify the registration standards for achievements, break information barriers, and achieve precise matching and efficient circulation of innovation resources.

3.4 Establish a Full-Cycle Financial Support System and Cultivate Patient Capital

To address the issues of capital maturity mismatch and risk aversion, a financial support system covering the entire life cycle of technological transformation needs to be established. In the seed and start-up stages, vigorous development of angel investment and venture capital should be pursued. Through tax incentives, risk compensation, government-led funds providing concessions, etc., guide social capital to invest in early-stage hard technology projects. Establish a national technology transfer guidance fund sub-fund and clearly stipulate the proportion of investment in "early-stage, small-scale, and hard technology" projects. In the growth stage, develop financial tools such as intellectual property pledge financing, technology insurance, and investment-lending linkage. Support commercial banks in establishing technology finance divisions or technology branches, and develop credit evaluation models and credit products suitable for technology enterprises. Explore the establishment of an intellectual property assessment and disposal mechanism to provide a foundation for pledge financing. In the mature stage, support qualified hard technology enterprises to utilize the multi-level capital market for financing. Improve the functional positioning of the Sci-Tech Innovation Board and the Growth Enterprise Market to serve technological transformation, and smooth the capital exit channels. Explore the establishment of an S fund (second-hand share transfer fund) to provide an exit channel for early-stage investment. By establishing a financial support system covering the entire life cycle of technological transformation, provide abundant "blood" for the cultivation of new productive forces.

4. Conclusion and Outlook

This study focuses on the improvement of the transformation efficiency of scientific and technological achievements under the background of new quality productivity, starts from the diagnosis of structural problems, deeply analyzes the deep causes of the problems, and systematically puts forward the policy optimization path. It is concluded that at present, the transformation system of scientific and technological achievements in China is faced with multiple obstacles, such as structural break of innovation chain, structural imbalance of subject relationship, and structural defect of service system. These structural problems are not accidental local obstacles, but the result of long-term accumulation of multiple factors

Based on the above analysis, this study proposes that the policy path to improve the efficiency of the transformation of scientific and technological achievements should adhere to the system thinking, from the subject activation, chain connection, ecological cultivation and other aspects of coordinated promotion: deepen the reform of property rights system and income distribution system, solve the dilemma of incentive incompatibility; Reshaping the innovative organizational model, and filling the "valley of death" through institutional design such as proof-of-concept centers and market-oriented pilot test platforms; Cultivate professional service ecology, accelerate the construction of technical managers and unified technical factor market construction; We will build a financial support system that covers the whole cycle of achievement transformation, and vigorously develop patient capital.

Facing the future, with the deepening of the concept of new quality productivity and the promotion of practice, there is still a huge space for the transformation of scientific and technological achievements. For example, digital technology, especially artificial intelligence, has a profound impact on the transformation mode of scientific and technological achievements.

The boundary reconstruction of government and market in the transformation of scientific and technological achievements under the new nationwide system, how to realize the organic combination of efficient market and effective government, etc. Continuous exploration of these issues will help deepen the understanding of the law of transformation of scientific and technological achievements, and provide more solid theoretical support for accelerating the formation of new quality productive forces and realizing high-quality development.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Song, D. Y., & Chen, L. (2024). Theoretical logic, key issues and practical path of developing new quality productivity. Review of Economy and Management, 40(05), 55–68.

[2] Ma, J. T., Cai, F., Gao, P. Y., et al. (2024). The talks on paper of the study and implement the spirit of the Third Plenary Session of the 20th Central Committee of the Communist Party of China. China Industrial Economics, (07), 5–28.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Measurement and Evaluation of Regional Competitiveness of the Nine Provinces in the Yellow River Basin Based on Entropy Weight-TOPSIS Method

Bingxin Wang*

School of Public Administration, Shandong Technology and Business University, Yantai, Shandong, 264005, China

Corresponding author: Bingxin Wang

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: The Yellow River Basin occupies an important strategic position in China's ecological security and economic development pattern. Scientifically measuring the regional competitiveness of the nine provinces in the basin and revealing its spatial differentiation characteristics are of great practical significance for promoting ecological protection and high-quality development of the Yellow River Basin. Based on the new development philosophy, this paper constructs a regional competitiveness evaluation system covering 22 specific indicators from seven dimensions: economic growth, industrial structure, cultural industry, green ecology, opening-up, infrastructure, and innovation capacity. The Entropy Weight-TOPSIS method is adopted to comprehensively measure and systematically analyze the competitiveness level of the nine provinces in the Yellow River Basin. The results show that the regional competitiveness of the nine provinces presents significant spatial disequilibrium, with a gradient distribution pattern of "strong in the east, weak in the west, and gradually declining from east to west". Shandong Province ranks first in comprehensive competitiveness, followed by Sichuan and Henan, while Qinghai and Gansu are at the bottom. From the dimensional perspective, obvious differences exist in economic growth, industrial structure and opening-up; midstream provinces perform better in green ecology; innovation capacity has become the core factor leading to competitiveness differentiation, and cultural industry is a common shortcoming across the basin. Overall, the competitiveness gap in the basin results from the combined effects of economic foundation, industrial structure, innovation factors and ecological background. Finally, policy implications are put forward from strengthening leading drive, implementing targeted policies, and focusing on innovation and green development, so as to provide a reference for synergistically improving the regional competitiveness of the nine provinces and realizing coordinated regional development in the Yellow River Basin.

Keywords: Yellow River Basin; Regional Competitiveness; Entropy Weight-TOPSIS Method; Spatial Differentiation; High-Quality Development

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1400

1. Introduction

The Yellow River Basin plays a vital role in China's economic and social development and ecological security [1]. The provinces it flows through are not only important components of high-quality economic development in the basin, but also a solid foundation for enhancing basin competitiveness, which is crucial for stabilizing the northern economy and promoting

economic construction with the domestic circulation as the mainstay [2]. Since the 18th National Congress of the Communist Party of China, the development and protection of the Yellow River Basin have received unprecedented attention, which has greatly boosted the regional competitiveness of the nine provinces in the basin. By 2021, the total population of the nine provinces in the Yellow River Basin reached 420 million, accounting for one-third of the national total, and their regional GDP amounted to nearly 30 trillion yuan, making up one-quarter of the national total. Although the nine provinces have made remarkable progress in development, there are still heterogeneities among provinces in geographical location, natural environment, resource endowment, talent scale, industrial policies, opening-up level and other aspects, leading to obvious gaps in regional competitiveness. To improve the regional competitiveness of each province in a more balanced, sustainable and stable manner, it is necessary to scientifically measure the regional competitiveness of the nine provinces, form a situation of complementary advantages and collaborative interaction among them, and thus promote coordinated regional development. This is of great significance for facilitating the implementation of the strategy for ecological protection and high-quality development of the Yellow River Basin.

Regional competition refers to competitive activities among multiple regions based on resource factors. Well-known academic discussions on international competitiveness include the studies conducted by the World Economic Forum (WEF), the International Institute for Management Development (IMD) and Michael Porter on international competitiveness. In 1989, the former State Commission for Restructuring the Economic System began to cooperate with WEF-IMD to study international competitiveness, and since then Chinese scholars have launched research on regional competitiveness. At present, domestic scholars mainly study regional competitiveness by constructing evaluation index systems, but most studies focus on the national, inter-provincial and urban levels. For example, Huang Guoyan, Liu Jianghui, et al. constructed an evaluation index system for global urban network competitiveness and compared the competitiveness of major cities at home and abroad [3]; Li Lin and Cao Can established an evaluation system for regional economic development in Hunan Province from six aspects: effectiveness, stability, innovation, coordination, greenness and sharing [4]; Dong Suocheng, Li Zehong, et al. constructed an urban competitiveness evaluation index system for state-level new areas from the perspectives of national development strategy, economic development and location conditions [5]; Zhang Chao, Li Changbiao, et al. built an evaluation system for regional coordinated development from five dimensions: economic development, public services, infrastructure, people's livelihood and ecological environment [6]; Wu Shaohua and Li Yujia constructed an urban competitiveness evaluation index system for western China from economic strength, infrastructure, social security and other aspects [7]; Xiao Feng, Wang Peng, et al. established a provincial competitiveness evaluation index system for the Yangtze River Economic Belt from economic growth and industrial structure [8].

Through literature review, it can be seen that the evaluation scope of regional competitiveness covers the national, inter-provincial and urban levels, but there are few studies taking the Yellow River Basin as the research object. The research methods mainly rely on constructing index evaluation systems to analyze regional competitiveness, and the research contents mostly focus on high-quality development [9] and ecological protection [10], lacking research on the regional competitiveness of provinces in the Yellow River Basin. Therefore, this paper takes the nine provinces in the Yellow River Basin as the research object, measures and analyzes their regional competitiveness, and explores the current development status and existing problems, so as to provide theoretical support and practical reference for the high-quality development of the Yellow River Basin.

2. Construction of the Evaluation Index System for Regional Competitiveness

2.1 Study Area

The Yellow River Basin crosses four major geomorphic units from west to east: the Qinghai-Tibet Plateau, the Inner Mongolia Plateau, the Loess Plateau, and the Huang-Huai-Hai Plain. It flows through nine provincial-level regions: Qinghai, Sichuan, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, Henan, and Shandong. With a vast basin area and a large population, the Yellow River Basin serves as a critical ecological barrier and economic zone in China. Meanwhile, the basin is characterized by a fragile ecological environment and a severe water resource supply situation. The ecological protection and high-quality development of the Yellow River Basin has become another national strategy following the Yangtze

River Economic Belt, the Yangtze River Delta Integration, the Beijing-Tianjin-Hebei Coordinated Development, and the Guangdong-Hong Kong-Macao Greater Bay Area.

2.2 Construction of the Evaluation Index System

The regional competitiveness evaluation index system provides a comprehensive assessment of economy, society, ecology, industry, culture, and other dimensions. In establishing the index system for the nine provinces in the Yellow River Basin, this paper adheres to the principles of hierarchy, operability, and comprehensiveness. With reference to relevant studies by Xiao Feng, Wang Peng et al.[11] and Ren Baoping, Gong Yuhao et al.[12], and in line with the new development philosophy featuring innovation, coordination, green development, opening-up, and sharing, this study takes full account of the economic development and cultural industry performance of each province. Seven first-level indicators are determined: economic growth, industrial structure, cultural industry, green ecology, opening-up, infrastructure, and innovation capacity. On this basis, 22 second-level indicators are selected, including 21 positive indicators and 1 negative indicator. The measurement index system of regional competitiveness of the nine provinces in the Yellow River Basin is shown in Table 1.

Table 1 Index System for Measuring Regional Competitiveness of the Nine Provinces in the Yellow River Basin

First-level Indicators	Second-level Indicators	Attribute	Weight
Economic Growth (Foundation)	Gross Regional Product (100 million yuan)	Positive	0.0463
	Fixed Asset Investment (100 million yuan)	Positive	0.0477
	Local Fiscal Revenue (100 million yuan)	Positive	0.0483
Industrial Structure (Guarantee)	Added Value of Secondary Industry (100 million yuan)	Positive	0.0353
	Added Value of Tertiary Industry (100 million yuan)	Positive	0.0446
	Total Electricity Consumption (100 million kWh)	Positive	0.0477
Cultural Industry (Performance)	Number of Cultural and Related Industries Units	Positive	0.0134
	Employees in Cultural and Related Industries	Positive	0.0403
	Operating Income of Cultural Industries (10,000 yuan)	Positive	0.036
Green Ecology (Requirement)	Forest Coverage Rate (%)	Positive	0.0514
	Sulfur Dioxide Emissions (10,000 tons)	Negative	0.0431
	Green Coverage Rate in Built-up Areas (%)	Positive	0.0555
Opening-up (Approach)	Foreign Investment (100 million yuan)	Positive	0.053
	Total Imports and Exports of Goods (10,000 yuan)	Positive	0.0291
	International Tourism Income (million USD)	Positive	0.0371
Infrastructure (Content)	Number of Urban Basic Medical Insurance Participants (10,000 persons)	Positive	0.0414
	Number of Doctors per 1,000 Persons	Positive	0.0455
	Number of Public Libraries	Positive	0.0451
	Per Capita Urban Road Area (sq.m.)	Positive	0.0538
Innovation Capacity (Momentum)	R&D Expenditure (100 million yuan)	Positive	0.0481
	Patent Applications	Positive	0.0419
	College Students per 100,000 Persons	Positive	0.0398

2.2.1 Economic Growth Index System

Economic growth is the solid foundation for improving regional competitiveness, reflecting the productivity and economic momentum of provinces, as well as a prerequisite for modernization and sustainable development in the Yellow River Basin. Indicators including gross regional product, fixed asset investment, and local fiscal revenue are selected to measure economic growth.

2.2.2 Industrial Structure Index System

Industrial structure is an important guarantee for enhancing regional competitiveness. A reasonable industrial structure helps form a sound industrial cycle and rational division of labor, narrowing gaps among upper, middle, and lower reaches. Given the long-term dominance of heavy and chemical industries in the Yellow River Basin, indicators such as added value of secondary and tertiary industries and total electricity consumption are used to analyze the rationalization level of industrial structure[13].

2.2.3 Cultural Industry Index System

As an important birthplace of traditional Chinese culture, the Yellow River Basin boasts unique advantages for cultural industry development. Three indicators—number of cultural enterprises, employees, and operating income—are adopted to measure input and output performance of the cultural industry.

2.2.4 Green Ecology Index System

Green development is a strategic requirement for the Yellow River Basin. Given its fragile ecology and severe pollution, forest coverage rate, sulfur dioxide emissions, and green coverage rate in built-up areas are selected to reflect ecological protection and environmental governance.

2.2.5 Opening-up Index System

Opening-up is an effective way to boost regional competitiveness. Foreign investment, total imports and exports, and international tourism income are used to measure the degree of opening-up and external connectivity.

2.2.6 Infrastructure Index System

Infrastructure supports high-quality development and regional connectivity. Indicators including medical insurance coverage, doctors per capita, public libraries, and per capita road area are selected to reflect infrastructure and public service levels.

2.2.7 Innovation Capacity Index System

Innovation is the long-term driving force for competitiveness. R&D expenditure, patent applications, and college students per 100,000 persons reflect innovation input, output, and talent reserve.

2.3 Data Sources and Research Methods

Statistical data are mainly obtained from China Statistical Yearbook (2023), China Fixed Investment Statistical Yearbook, and China Fixed Investment Statistical Annual Report. Missing 2023 data are supplemented by 2022 provincial data.

The Entropy Weight-TOPSIS method is adopted for comprehensive evaluation. This method combines entropy weight (for objective weighting) and TOPSIS (for ranking by relative closeness to the ideal solution). The comprehensive score C is calculated; a value closer to 1 indicates stronger competitiveness.

3. Analysis of Evaluation Results of Regional Competitiveness

3.1 Model Results

Indicator weights are calculated using the entropy weight method, as shown in Table 2.

Table 2 Weight Results Calculated by Entropy Weight Method

Secondary Indicators	Entropy Value e	Information Utility Value d	Weight Coefficient w
Gross Regional Product (100 million yuan)	3.884992	-2.88499169	0.0463471
Fixed Asset Investment (100 million yuan)	3.968548	-2.96854811	0.0476894
Local Fiscal Revenue (100 million yuan)	4.008212	-3.00821175	0.0483266
Added Value of Secondary Industry (100 million yuan)	3.198202	-2.19820227	0.0353139

Secondary Indicators	Entropy Value e	Information Utility Value d	Weight Coefficient w
Added Value of Tertiary Industry (100 million yuan)	3.775001	-2.77500068	0.0445801
Total Social Electricity Consumption (100 million kWh)	3.9677	-2.96769969	0.0476758
Number of Cultural and Related Industry Units (units)	1.83363	-0.83363035	0.0133922
Employees in Cultural and Related Industries (persons)	3.506054	-2.50605403	0.0402595
Operating Income of Cultural and Related Industries (10,000 yuan)	3.243658	-2.24365783	0.0360442
Forest Coverage Rate (%)	4.196501	-3.19650096	0.0513515
Per Capita Green Space Area (square meters)	3.684057	-2.68405653	0.0431191
Sulfur Dioxide Emissions (10,000 tons)	4.454085	-3.45408503	0.0554896
Green Coverage Rate in Built-up Areas (%)	4.298366	-3.29836589	0.0529879
Foreign Investment Amount (100 million yuan)	2.813399	-1.81339873	0.0291321
Total Import and Export Volume of Foreign Trade Goods (10,000 yuan)	3.308228	-2.30822849	0.0370815
International Tourism Income (million USD)	3.575699	-2.5756987	0.0413784
Number of Urban Basic Medical Insurance Participants (10,000 persons)	3.830542	-2.83054205	0.0454724
Number of Doctors per 1,000 Persons (persons)	3.80629	-2.80628981	0.0450828
Number of Public Libraries (units)	4.346279	-3.34627943	0.0537577
Per Capita Urban Road Area (square meters)	3.997122	-2.9971217	0.0481485
R&D Expenditure (100 million yuan)	3.606185	-2.60618486	0.0418681
Number of Patent Applications (pieces)	3.478879	-2.47887915	0.039823
Number of College Students per 100,000 Persons (persons)	4.465859	-3.465859	0.0556787

The positive ideal solution (R^+) and negative ideal solution (R^-) are obtained, and Euclidean distances D^+ and D^- are calculated (Table 3). The comprehensive closeness coefficient C is then derived and ranked (Table 4).

Table 3 Positive and Negative Ideal Solutions

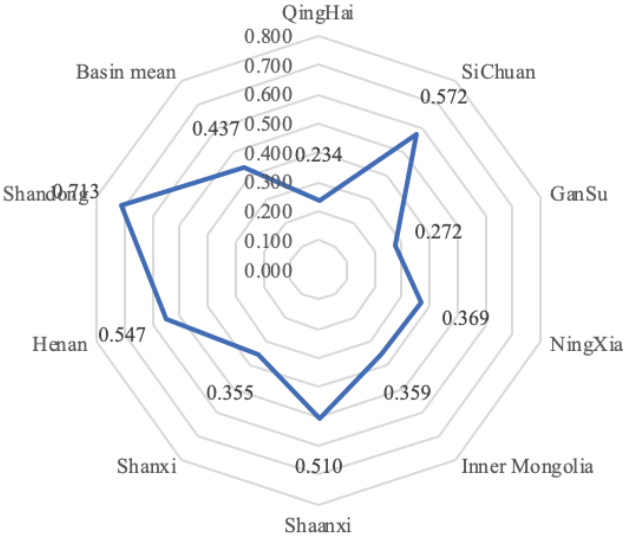
Secondary Indicators	Positive Ideal Solution R+	Negative Ideal Solution R−
Gross Regional Product (100 million yuan)	0.046393	0.000046
Fixed Asset Investment (100 million yuan)	0.047737	0.000048
Local Fiscal Revenue (100 million yuan)	0.048375	0.000048
Added Value of Secondary Industry (100 million yuan)	0.035349	0.000035
Added Value of Tertiary Industry (100 million yuan)	0.044625	0.000045
Total Social Electricity Consumption (100 million kWh)	0.047723	0.000048
Number of Cultural and Related Industry Units (units)	0.013406	0.000013
Employees in Cultural and Related Industries (persons)	0.0403	0.00004
Operating Income of Cultural and Related Industries (10,000 yuan)	0.03608	0.000036
Forest Coverage Rate (%)	0.051403	0.000051
Per Capita Green Space Area (square meters)	0.043162	0.000043
Sulfur Dioxide Emissions (10,000 tons)	0.055545	0.000055
Green Coverage Rate in Built-up Areas (%)	0.053041	0.000053

Secondary Indicators	Positive Ideal Solution R+	Negative Ideal Solution R−
Foreign Investment Amount (100 million yuan)	0.029161	0.000029
Total Import and Export Volume of Foreign Trade Goods (10,000 yuan)	0.037119	0.000037
International Tourism Income (million USD)	0.04142	0.000041
Number of Urban Basic Medical Insurance Participants (10,000 persons)	0.045518	0.000045
Number of Doctors per 1,000 Persons (persons)	0.045128	0.000045
Number of Public Libraries (units)	0.053811	0.000054
Per Capita Urban Road Area (square meters)	0.048197	0.000048
R&D Expenditure (100 million yuan)	0.04191	0.000042
Number of Patent Applications (pieces)	0.039863	0.00004
Number of College Students per 100,000 Persons (persons)	0.055734	0.000056

Table 4 Relative Closeness C and Ranking of Regional Competitiveness of Provinces in the Yellow River Basin

Province	Relative Closeness C	Ranking
Qinghai	0.234	9
Sichuan	0.572	2
Gansu	0.272	8
Ningxia	0.369	5
Inner Mongolia	0.359	6
Shaanxi	0.51	4
Shanxi	0.355	7
Henan	0.547	3
Shandong	0.713	1

Figure 1 Relative Closeness C of Regional Competitiveness Level of Provinces in the Yellow River Basin



3.2 Analysis of the Evaluation and Measurement Results of Regional Competitiveness of the Nine Provinces in the Yellow River Basin

From the measurement results in Table 4 and the comprehensive competitiveness radar chart, the regional competitiveness

of the nine provinces in the Yellow River Basin presents significant spatial disequilibrium and a stepped gradient distribution feature. The overall pattern follows the law of "gradual attenuation from east to west", which is highly consistent with the economic and geographical location, resource endowment conditions and policy implementation intensity in the basin.

In terms of comprehensive competitiveness, Shandong Province ranks first with a relative closeness value of 0.713, becoming the absolute leading engine for high-quality development in the Yellow River Basin. Its scores on all indicators far exceed the average level of the basin, showing a gap-leading advantage in the comprehensive competitiveness dimension, and it is the only first-echelon province with all-round comprehensive strength in the basin. Sichuan Province (0.572) and Henan Province (0.547), which follow closely, rank second and third respectively. As the southwest growth pole and the core of the Central Plains hinterland, they have constructed a dual-core-driven growth pattern of "east, central and west" in the Yellow River Basin. Shaanxi Province (0.510) ranks fourth by virtue of its advantages in science, education and technological innovation, and is in the transition zone from the first echelon to the second echelon.

Ningxia Hui Autonomous Region (0.369), Inner Mongolia Autonomous Region (0.359) and Shanxi Province (0.355), which rank in the middle reaches of the comprehensive competitiveness, have relatively close C values, forming a distinctive intermediate force. The provinces in this echelon generally have a resource-based industrial foundation or specific location advantages, but there are still shortcomings in the balanced allocation of comprehensive factors, facing the dual pressures of industrial transformation and ecological governance, and their competitiveness level is basically above the average level of the basin.

Qinghai Province (0.234) and Gansu Province (0.272), which rank at the end of the basin, have comprehensive competitiveness scores significantly lower than the basin average, with a huge gap from the leading provinces. In particular, despite its unique ecological background, Qinghai Province is restricted by factors such as small economic volume, backward infrastructure and lack of innovation factors, and its comprehensive competitiveness is in the weakest position in the whole basin.

On the whole, the regional competitiveness of the Yellow River Basin presents a dumbbell-shaped structure of "high concentration in the head, differentiated echelons in the waist, and overall lag in the tail". This pattern not only reflects the objective gaps in economic aggregate, public services and innovation capacity between the eastern, central and western regions, but also reveals the practical challenges in the process of coordinated development of the Yellow River Basin, such as how to narrow the gap between the east and the west through policy inclination and achieve a win-win situation between ecological protection and economic growth. The leading advantages of Shandong, Henan and Sichuan are not only reflected in economic scale, but also in the comprehensive control ability of total factor productivity; while the backwardness of upstream provinces is mainly restricted by structural obstacles such as ecological burden, weak foundation and brain drain.

To investigate the dimensional differences among the nine provinces in the Yellow River Basin, a further comparison is made on the dimensional measurement values of the nine provinces in 2022. The comparisons of regional competitiveness in each dimension are shown in Figures 2–8.

Figure 2 Economic Growth Dimension

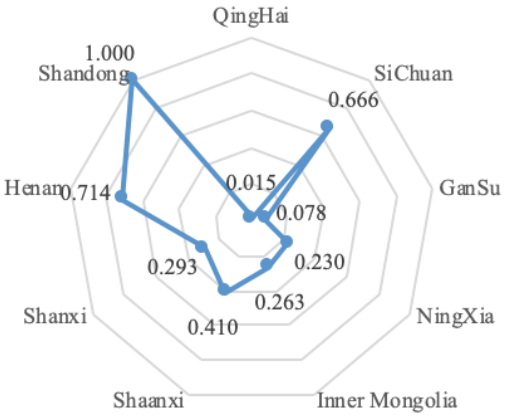


Figure 3 Industrial Structure Dimension

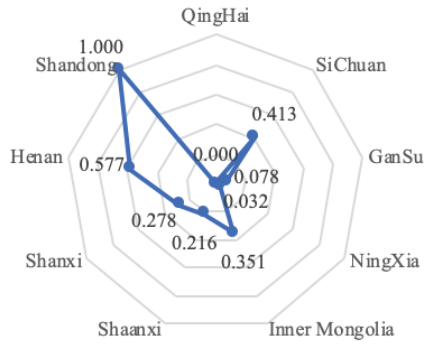


Figure 4 Cultural Industry Dimension

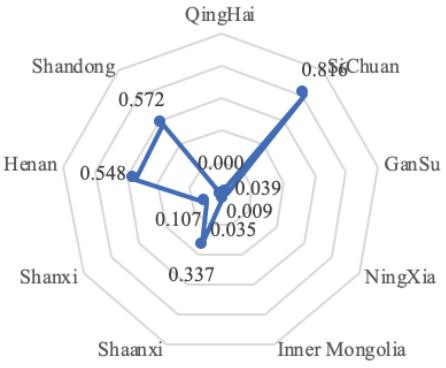


Figure 5 Green Ecology Dimension

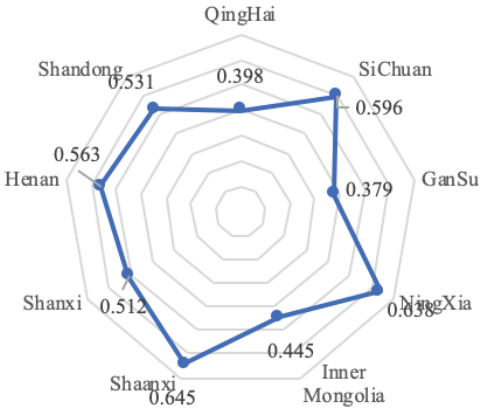


Figure 6 Opening-up Dimension

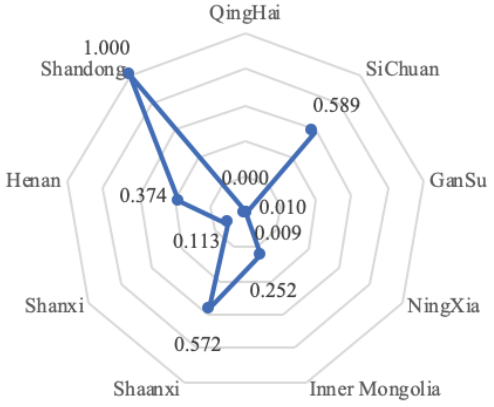


Figure 7 Infrastructure Dimension

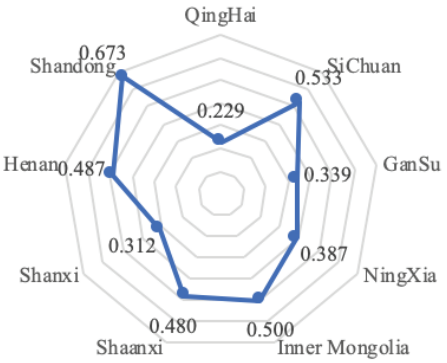
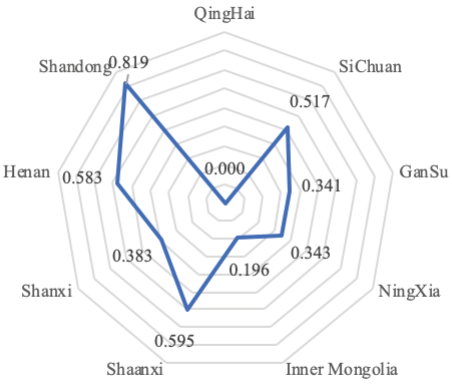


Figure 8 Innovation Capacity Dimension



3.2.1 Economic Growth Dimension

According to the evaluation results of economic growth, the nine provinces in the Yellow River Basin present an extremely significant gradient differentiation pattern. Shandong takes a leading position with a full score of 1.000, followed by Henan (0.7136) and Sichuan (0.6665), which together constitute the first echelon of economic growth in the basin. Shaanxi, Shanxi, Inner Mongolia and Ningxia are at the middle level with scores ranging from 0.2 to 0.4. Gansu and Qinghai score only 0.0783 and 0.0154 respectively, over 60 times lower than the leading provinces, placing them in an absolutely weak position. Such a differentiated pattern stems from the combined effects of location, population size, industrial structure and fiscal capacity. Shandong and Henan are located in the lower Yellow River plain with large populations and high urbanization rates, possessing inherent advantages for economic expansion. Their core indicators such as GDP, local fiscal revenue and fixed asset investment have long ranked top in the basin, making them traditional growth poles. Sichuan, relying on the strategic dividend of the Chengdu-Chongqing Economic Circle, has become a core hub of population and industry aggregation in southwest China, surpassing midstream resource-based provinces in economic scale and forming a dual-core growth pattern of the lower reaches and southwest China. In contrast, Shaanxi, Shanxi and Inner Mongolia in the middle reaches take energy and heavy chemical industries as the leading sectors, with economic growth highly dependent on resource price fluctuations. Their monotonous industrial structure leads to insufficient growth momentum and weak risk resistance. Gansu and Qinghai in the upper reaches are restricted by small economic aggregates and low fiscal self-sufficiency, with weak driving effects of fixed asset investment, falling into a vicious circle of weak economy, low investment and slow growth. This eventually forms an extreme pattern of leadership in the lower reaches and collapse in the upper reaches, which constitutes the fundamental cause of the overall imbalance of regional competitiveness in the Yellow River Basin.

3.2.2 Industrial Structure Dimension

The evaluation results of industrial structure maintain the overall trend of strong east and weak west. Shandong takes an absolute lead with a full score of 1.000, and Henan (0.577) and Sichuan (0.413) rank second and third respectively,

forming the first echelon of industrial structure optimization. Inner Mongolia, Shanxi and Shaanxi are in the middle stage of transformation with scores between 0.2 and 0.35. Gansu, Ningxia and Qinghai score less than 0.1, and Qinghai ranks last with 0.000, showing prominent shortcomings in industrial structure. Fundamentally, this difference reflects the comprehensiveness of industrial systems, the advancement of structure and the effectiveness of transformation. As one of the provinces with the most complete industrial categories in China, Shandong boasts coordinated development of secondary and tertiary industries, driven by both manufacturing and modern services, leading the basin in rationalization and advancement of industrial structure. Relying on the geographical advantage of the Central Plains Urban Agglomeration, Henan has built a complete industrial chain dominated by equipment manufacturing and food processing, with its industrial structure continuously upgraded to become a benchmark for structural optimization in the middle and lower reaches. However, as traditional energy-intensive provinces, Inner Mongolia, Shanxi and Shaanxi have long maintained a high proportion of secondary industry with a heavy and monotonous structure and lagging service industries. Despite high total social electricity consumption, most is consumed by energy-intensive industries with low added value, and slow transformation restricts the improvement of structural competitiveness. The upper reaches including Gansu, Ningxia and Qinghai are dominated by agriculture, animal husbandry and primary energy processing with weak industrial foundations. Their tertiary industries rely on traditional consumer services with insufficient support from emerging industries, resulting in weak driving effects on economic growth. The final pattern of leading head, transforming middle reaches and collapsing upper reaches confirms that industrial structure optimization is the core approach to enhancing regional competitiveness.

3.2.3 Cultural Industry Dimension

The evaluation of cultural industry shows a distinctive pattern of Sichuan standing out and the whole basin lagging behind generally. Sichuan takes a leading position with a high score of 0.816, followed by Shandong (0.572) and Henan (0.548) as the first echelon. Shaanxi and Shanxi score less than 0.4 at the middle level, while Gansu, Inner Mongolia, Ningxia and Qinghai all score below 0.04, with Qinghai ranking bottom at 0.000, nearly devoid of cultural industry competitiveness. The core cause lies in the serious mismatch between cultural resource endowment and industrial transformation capacity. As the core birthplace of Chinese civilization, the nine provinces boast abundant historical, revolutionary and intangible cultural heritage, yet "large in resources but small in industry" is a common problem across the basin. Relying on the culture-tourism integration strategy of the Chengdu-Chongqing Economic Circle, Sichuan has successfully developed world-class cultural IPs such as Sanxingdui and Jiuzhaigou, ranking first in the basin in the number of cultural enterprises, employees and business income, realizing efficient transformation from cultural resources to industrial competitiveness. Shandong and Henan have promoted large-scale development of cultural industries by virtue of population and market advantages, forming a certain industrial foundation. Nevertheless, Shaanxi and Shanxi, despite possessing top cultural resources such as Qin and Jin cultures, suffer from low industrialization, lack of nationally influential cultural IPs and leading enterprises, failing to convert resource advantages into industrial strengths. Restricted by economic scale, market capacity and talent reserve, the upper-reach provinces including Gansu, Qinghai, Ningxia and Inner Mongolia have a small number of cultural market entities and low business income with insufficient market-oriented operation. Cultural resource development mostly stays at the primary tourism stage without forming a complete industrial chain, making cultural industry a common shortcoming of competitiveness in the basin and one of the areas with the greatest potential for future improvement.

3.2.4 Green Ecology Dimension

The green ecology evaluation breaks the traditional pattern of strong east and weak west, featuring leadership in the middle reaches, follow-up in the upper reaches and mediocrity in the lower reaches. Shaanxi and Ningxia rank top with scores of 0.645 and 0.638 respectively. Sichuan, Henan, Shandong and Shanxi are at the upper-middle level with scores between 0.5 and 0.6. Inner Mongolia, Qinghai and Gansu rank downstream with scores from 0.37 to 0.45. This pattern results from ecological background, pollution control and urban greening. As core regions for ecological protection in the middle Yellow River, Shaanxi and Ningxia have continuously implemented ecological projects such as returning farmland to forest and desert control in recent years, performing well in green coverage rate of built-up areas. Meanwhile, they have strictly controlled sulfur dioxide emissions, realizing coordinated ecological protection and pollution control. Sichuan and Qinghai,

relying on the ecological advantages of the upper Yangtze and Yellow Rivers, rank high in natural ecological indicators such as forest coverage with obvious foundational strengths. As densely populated and industrialized lower-reach provinces, Shandong and Henan have achieved remarkable results in pollution reduction and urban greening despite weaker ecological backgrounds, balancing economic development and ecological protection. As traditional energy provinces, Shanxi and Inner Mongolia have gradually improved ecological competitiveness through energy restructuring and pollution control despite high sulfur dioxide emissions. Gansu, restricted by ecological fragility and economic pressure, faces great difficulties in balancing protection and governance, thus ranking low. Overall, the green ecology results confirm that "ecological priority and green development" has become a core criterion for regional competitiveness in the Yellow River Basin, where pollution reduction and ecological restoration directly determine regional ecological competitiveness.

3.2.5 Opening-up Dimension

The opening-up evaluation presents an extremely differentiated pattern of absolute leadership by Shandong, follow-up by southwest provinces and collapse in the upper reaches. Shandong takes a leading position with a full score of 1.000, followed by Sichuan (0.589) and Shaanxi (0.572), and Henan ranks fourth with 0.374. Inner Mongolia and Shanxi score less than 0.3, while Gansu, Ningxia and Qinghai score below 0.01, with Qinghai at the bottom of 0.000, showing a huge gap in opening-up levels. The core cause lies in differences in geographical conditions, opening channels and export-oriented industrial foundations. As the only coastal province in the Yellow River Basin, Shandong has world-class ports such as Qingdao and Yantai, ranking first in core indicators including foreign trade volume, foreign investment and international tourism income, serving as the leading gateway for opening-up. Sichuan and Shaanxi, relying on opening channels such as China-Europe Railway Express and international aviation hubs, have built highlands for inland opening-up with sound export-oriented industries and high opening levels. Henan, supported by Zhengzhou Airport and China-Europe Railway Express (Zhengzhou), has become an important opening node in the middle and lower reaches. However, Inner Mongolia and Shanxi, despite border-opening advantages, suffer from weak export-oriented industries and small foreign trade scales. The inland upper-reach provinces including Gansu, Qinghai and Ningxia lack sufficient opening channels and lag in export-oriented industries, with minimal foreign capital utilization, foreign trade and international tourism income, resulting in extremely low opening levels and becoming a core bottleneck restricting the improvement of regional competitiveness. Overall, the opening-up level of the Yellow River Basin is generally low with insufficient release of opening dividends, and the opening shortcomings of inland provinces are key breakthroughs for coordinated development.

3.2.6 Infrastructure Dimension

The infrastructure evaluation shows a pattern of leadership by Shandong, mediocrity of multiple provinces and weakness in the upper reaches. Shandong ranks first with a high score of 0.673. Sichuan, Inner Mongolia, Henan and Shaanxi are at the upper-middle level with scores from 0.48 to 0.54. Ningxia, Gansu and Shanxi score between 0.3 and 0.4, and Qinghai ranks last with 0.229. This pattern is jointly shaped by economic strength, population size and policy investment. As an economically strong province, Shandong boasts strong fiscal capacity, leading comprehensively in infrastructure and public service indicators such as urban basic medical insurance participants, doctors per 1,000 people, public libraries and per capita urban road area, with the strongest infrastructure support capacity in the basin. As large population and economic provinces, Sichuan and Henan invest heavily in infrastructure with complete public service systems and high infrastructure levels. Inner Mongolia and Shaanxi, supported by fiscal revenue from energy industries, have continuously improved infrastructure to a medium-upper level. Restricted by economic scale and fiscal strength, Ningxia, Gansu and Shanxi invest insufficiently in infrastructure with low equalization of public services. Qinghai, due to sparse population and complex geographical conditions, faces high costs and difficulties in infrastructure construction with low per capita infrastructure levels, thus ranking bottom. Overall, infrastructure competitiveness is highly positively correlated with regional economic strength. Infrastructure perfection directly determines regional development carrying capacity and livelihood security, and remedying infrastructure shortcomings in upper-reach provinces is crucial for narrowing regional gaps.

3.2.7 Innovation Capacity Dimension

The innovation capacity evaluation presents a pattern of leadership by Shandong, follow-up by Shaanxi, Henan and Sichuan,

and lagging in the upper reaches. Shandong takes a leading position with a high score of 0.819. Shaanxi (0.595), Henan (0.583) and Sichuan (0.517) rank second to fourth respectively, forming the first echelon of innovation capacity. Shanxi, Ningxia and Gansu are at the middle level with scores from 0.34 to 0.39. Inner Mongolia and Qinghai score less than 0.2, and Qinghai ranks last with 0.000, showing a significant innovation gap. The core cause lies in differences in scientific and educational resources, R&D investment and talent reserve. As an economically and educationally strong province, Shandong ranks first in core innovation indicators such as R&D expenditure, patent applications and college students per 100,000 people, with outstanding innovation-driven development capacity. Relying on science and education centers such as Xi'an and Chengdu, Shaanxi and Sichuan boast abundant university resources, high R&D intensity and strong patent output, becoming core growth poles for innovation in the basin. Henan, supported by population and industrial advantages, has continuously increased R&D investment with steadily improving innovation capacity. As resource-based provinces, Shanxi and Inner Mongolia suffer from insufficient R&D investment and shortage of high-end talents with lagging innovation capacity. Restricted by shortage of scientific and educational resources and serious brain drain, the upper-reach provinces including Gansu, Qinghai and Ningxia have weak ability to gather innovation factors, minimal R&D investment and patent output, nearly devoid of innovation-driven capacity, thus ranking bottom. Overall, innovation capacity is highly positively correlated with comprehensive regional competitiveness. Talent and technology have become core variables differentiating regional competitiveness in the Yellow River Basin, and innovation-driven development is the fundamental momentum for high-quality development.

4. Conclusions and Policy Recommendations

This study comprehensively measures the regional competitiveness of the nine provinces in the Yellow River Basin using the entropy weight-TOPSIS method. The results show that the regional competitiveness presents significant spatial disequilibrium and stepped gradient distribution, following the law of "strong in the east, weak in the west and gradual attenuation from east to west".

Comprehensively, Shandong ranks first in the basin with a high score of 0.713 as the only leading province with all-round advantages. Sichuan and Henan rank second and third, forming a dual-core growth pole of "southwest-central plains". Shaanxi ranks fourth. Ningxia, Inner Mongolia and Shanxi constitute the middle echelon. Qinghai and Gansu rank at the bottom with a huge competitiveness gap.

Dimensional analysis shows that economic growth, industrial structure and opening-up are traditional foundations widening the gap. Innovation capacity and green ecology have become key variables determining competitiveness differentiation, where regional gaps in scientific and educational resources and pollution control effectiveness directly decide competitiveness levels. As a common shortcoming across the basin, cultural industry reflects insufficient transformation of Yellow River cultural resources into industrial competitiveness. Overall, the competitiveness gap in the basin results from the combined effects of economic foundation, industrial structure, innovation factors and ecological background.

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

First, strengthen leading drive and multi-polar coordination to build a new pattern of coordinated basin development. Give full play to the radiation and driving role of Shandong as the basin leader, deepen its benchmark effect in opening-up, industrial upgrading and innovation-driven development, and drive the coordinated development of middle and lower-reach provinces. Rely on Sichuan and Henan to build a "southwest-central plains" dual-core growth pole, radiating the upper, middle and lower reaches respectively to form a development framework of leading drive, multi-polar support and regional linkage. Meanwhile, establish a cross-regional coordinated development mechanism for the Yellow River Basin, break administrative barriers, promote the free flow and efficient allocation of talents, capital, technology and other factors, and gradually narrow the development gap between east and west.

Second, implement targeted and precise policies to remedy dimensional development shortcomings. For midstream resource-based provinces, focus on promoting industrial structure transformation and upgrading, foster strategic emerging industries and modern services, and get rid of resource dependence. For underdeveloped upper-reach provinces, increase central fiscal transfer payments and policy support, remedy shortcomings in transportation, medical care, public culture

and other infrastructure, build a solid ecological security barrier, and foster green industries such as ecotourism and characteristic agriculture. For the common shortcoming of cultural industry, deeply explore the resources of Yellow River culture, revolutionary culture and intangible cultural heritage, build a world-class Yellow River cultural tourism belt, foster leading cultural enterprises and characteristic IPs, and promote the transformation of cultural resources into industrial competitiveness.

Third, anchor innovation-driven and green development as the core to activate long-term development momentum. Take innovation capacity as the core engine for improving regional competitiveness, coordinate the layout of scientific and educational resources in the basin, establish cross-regional talent sharing and joint R&D mechanisms, increase R&D investment in upper-reach provinces, and foster new innovation-driven momentum. Coordinately promote ecological protection and pollution control across the basin, improve the ecological compensation mechanism, promote green and low-carbon transformation of traditional energy provinces, realize the transformation of ecological benefits into economic benefits, and build a solid ecological foundation and inject sustained momentum into the high-quality development of the Yellow River Basin.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Xue, L., Yang, Y., Chen, L., et al. (2020). Legislative strategies for the strategy of ecological protection and high-quality development of the Yellow River Basin. China Population, Resources and Environment, 30(12).

[2] Jin, F. J., Lin, Y. H., Ma, L., et al. (2022). Evolution of strategic status and high-quality development direction of the Yellow River Basin. Journal of Lanzhou University (Social Sciences Edition), 50(1).

[3] Huang, G. Y., Liu, J. H., Jiang, X. T., et al. (2019). Comparative study on competitiveness of major cities at home and abroad based on complex network analysis. Urban Development Studies, 26(1).

[4] Li, L., & Cao, P. (2019). Quality evaluation of regional economic development in Hunan Province based on projection pursuit method. Journal of Hunan University (Social Sciences Edition), 33(3).

[5] Dong, S. C., Li, Z. H., Shi, G. Y., et al. (2022). Urban competitiveness evaluation of western prefecture-level cities and layout of state-level new areas in the 14th Five-Year Plan period. Journal of Natural Resources, 37(7).

[6] Zhang, C., Zhong, C. B., Jiang, T. Y., et al. (2020). Spatio-temporal differentiation of China's regional coordinated development and its influencing factors. Economic Geography, 40(9).

[7] Wu, S. H., & Li, Y. J. (2021). Study on urban competitiveness evaluation in western China based on principal component analysis. Economic Issues, (11).

[8] Xiao, F., Wang, P., & Chen, G. S. (2022). Provincial competitiveness evaluation of the Yangtze River Economic Belt and influencing factors of regional differences based on TOPSIS method. Economic Geography, (2).

[9] An, S. W., & Li, R. P. (2020). Connotation and promotion strategies of high-quality development of the Yellow River Basin. Reform, (1).

[10] Ren, B. P., & Gong, Y. H. (2022). Coupling study of urbanization and high-quality development in the Yellow River Basin. Economic Issues, (3).

[11] Ren, B. P., Fu, Y. M., & Yang, Y. C. (2022). Evaluation and path selection of high-quality economic development in nine provinces of the Yellow River Basin. Journal of Statistics and Information Forum, 37(1).

[12] Cui, G. R. (2022). Measurement and evaluation of high-quality development level of China's industry. Statistics & Decision, 38(5).

[13] Qiu, Z. L. (2022). Differences in ecological efficiency between the Yangtze River Economic Belt and the Yellow River Basin and their influencing factors. Contemporary Economic Management, (3).

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

ESG Integration and Long-Term Portfolio Performance: Evidence from International Markets

Jiale Cai*

School of Finance, Shandong Technology and Business University, Yantai, Shandong, 264005, China

Corresponding author: Jiale Cai, 663917068@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: In the macro context of the deep transformation of the global economy to a sustainable development model, environmental, social and corporate governance (ESG) factors are increasingly established in the core position of asset pricing and portfolio construction in the international capital market. This article explores the core mechanism of ESG integration on long-term portfolio performance and international market performance. The study shows that ESG integration is not a simple moral constraint, but significantly improves the long-term risk-adjusted returns of the portfolio by avoiding tail risks and optimizing the cost of capital of enterprises. However, due to the differences in market development stages, the excess returns of ESG investment show significant heterogeneity between developed and emerging markets, and face empirical challenges such as divergent rating standards and "greenwashing". Clarifying the above complex mapping mechanism provides a solid theoretical basis for institutional investors to optimize the cross-cycle asset allocation framework and regulatory authorities to improve the information disclosure system.

Keywords: ESG Integration; Portfolio Performance; Long-Term Gains; Excess Returns; Risk Management

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1405

1. Introduction

In the macro context of the global economy's deep transformation towards sustainable development, environmental, social and corporate governance (ESG) factors are profoundly reshaping the asset pricing logic of the international capital market. Traditional investment frameworks with a single financial indicator as the core are no longer able to effectively address systemic risks caused by climate change, supply chain restructuring, and corporate governance deficiencies. While mainstream institutions have embraced ESG concepts at a strategic level, controversy persists over whether they can generate robust excess returns for portfolios over the long term. While traditional investment theories worry that non-financial constraints will sacrifice diversified returns, the empirical evidence from a long-term perspective emphasizes that companies with high ESG performance tend to have lower financing costs, a more forward-looking strategic vision, and a stronger resilience to risks [1]. Therefore, relying on a large number of samples in the international market and deeply deconstructing the internal correlation between ESG integration and long-term portfolio performance, it is not only the frontier of the expansion of modern asset pricing theory, but also the key proposition to guide the green transformation of global capital.

2. The Mechanism of ESG Integration on Long-Term Portfolio Performance

To explore the underlying drivers of ESG integration on long-term portfolio performance, we must penetrate the surface

phenomenon and explore the bridge it builds between micro corporate financial characteristics and macro market pricing. ESG factors are not only a scale to measure corporate externalities, but also an internal engine that drives long-term capital appreciation.

2.1 Tail Risk Management based on the Perspective of Risk Mitigation

In today's uncertain business environment, risk management capabilities are a decisive factor in determining the long-term survival and prosperity of a portfolio. The primary mechanism of ESG integration at the portfolio level is reflected in the effective avoidance of "asymmetric downside risk". From a mechanical point of view, seamlessly embedding the ESG assessment system into the investment due diligence process can greatly broaden the breadth and depth of information acquisition, and help investors identify and accurately eliminate "black swan" candidates with fatal hidden dangers in environmental compliance, labor rights protection or corporate governance structure in advance. When the macroeconomy suffers a severe shock or an industry suddenly crisis breaks out, companies with high ESG ratings often show stronger resilience [2]. For example, in the face of increasingly stringent global carbon emission quota constraints and environmental penalties, enterprises that have made forward-looking layouts in the fields of green technology research and development and energy conservation and emission reduction in the early stage can calmly avoid huge compliance costs and the risk of production shutdown and rectification, and maintain the stability of free cash flow. This strong immunity to tail risk significantly reduces the volatility and maximum drawdown of the portfolio during the long-term investment cycle, thus maintaining the bottom line of long-term profitability for investors under the calculation rule of "compound interest magic".

2.2 Cost of Capital Optimization based on a Value-Driven Perspective

In addition to passive risk defense, ESG integration has an active value creation function, and its core transmission path is to significantly optimize the capital cost structure of enterprises. In the debt financing market, commercial banks and fixed income investors are increasingly inclined to include corporate ESG risk exposure in credit default probability measurement models. Companies that actively practice ESG concepts and have high transparency in information disclosure are often seen as having more stable operating expectations and stronger debt solvencies, allowing them to issue green bonds or obtain sustainability-linked loans at lower coupon rates. At the level of equity financing, excellent ESG performance not only sends a positive signal to the market that management is focused on long-term and stable operation, but also effectively reduces agency costs and information asymmetry between insiders and external investors. With the explosive growth of passive funds and actively managed funds that follow ESG investment principles around the world, companies with high ESG ratings can often enjoy significant "liquidity premiums" and higher valuation centers. This systematic reduction in the weighted average cost of capital (WACC) directly increases the net present value of the company's future discounted cash flow, which in turn drives the long-term stable growth of the intrinsic value of the portfolio.

2.3 Long-term Dynamic Integration of Financial and Non-Financial Performance

The improvement of long-term portfolio performance is not achieved overnight, but is the result of the dynamic integration of financial and non-financial elements over a long period of time. Traditional financial indicators are often lagging "rearview mirrors" that can only reflect the past operating performance of enterprises, while ESG indicators are strongly forward-looking and are the "telescope" for predicting the sustainable profitability of enterprises in the future. Empirical research has found that ESG factors have a unique "risk premium" characteristic, and this premium is more significant in the long-term perspective [3]. In addition, the widespread "ESG momentum effect" in the capital market provides a rich ground for actively managed portfolios to obtain excess returns. Static high-ESG rating companies may have been fully priced by the market and lack room for excess returns. However, those "in transition" companies that are in a period of ESG governance structure reshaping and whose indicators are showing a marginal improvement trend are often accompanied by huge dividends from fundamental revaluation. By prospectively capturing this non-financial marginal improvement at the micro level, institutional investors can lay out in advance and share the long-term growth dividends released by enterprises from "low expectations" to "high quality", so as to achieve deep resonance between financial performance and non-financial performance [4].

3. Heterogeneity and Strategic Evolution of ESG Investment Performance in the International Market

The international capital market is not a single homogeneous whole, and its internal differences in institutional soil, legal tradition and market effectiveness have shaped the profound heterogeneity of ESG integration effects and continue to promote the iterative upgrading of investment strategies.

3.1 Differences in Performance between Developed and Emerging Markets

Developed capital markets represented by Europe and North America are the source and mature testing ground of ESG investment concepts. In particular, the European market has built the world's most stringent and complete ESG information disclosure and evaluation infrastructure under the pressure of strong regulations such as the Sustainable Finance Disclosure Regulation (SFDR). In such a highly effective and crowded market environment, ESG information has been quickly reflected in asset prices by massive quantitative models and high-frequency trading funds, and it is exponentially more difficult to build a portfolio to obtain alpha returns solely based on public static ESG rating data. Investors' excess returns in developed markets increasingly rely on in-depth mining of unstructured data and the construction of complex nonlinear models. In stark contrast to the vast number of emerging markets, which are often accompanied by lagging disclosure systems, weak investor protection mechanisms, and pervasive corporate governance deficiencies. However, it is this data vacuum and market failure that provides a vast blue ocean for long-term funds that deeply practice ESG integration to obtain amazing excess returns. In emerging markets, the weight and influence of the "G" (corporate governance) factor often show an overwhelming advantage. By focusing on screening those companies with reasonable equity structure, norms of conduct of major shareholders, and transparent related party transactions, investors can avoid wealth plunder caused by agent moral hazard to the greatest extent. At the same time, with the restructuring of global supply chains, the harsh ESG access standards put forward by multinational giants for suppliers in emerging markets are forcing these companies to accelerate their transformation, thus giving birth to rich "ESG momentum" investment targets.

3.2 The Evolution of Investment Strategy from Negative Elimination to Comprehensive Integration

With the deepening of the market's understanding of the internal logic of sustainable development, the ESG strategies of international institutional investors have undergone a profound transformation from simple and crude to refined and deep. Early socially responsible investing mainly adopted a single "negative exclusion method", that is, based on specific religious beliefs or moral standards, mechanically excluding industries involved in the military, tobacco, gaming, or highly polluting fossil energy from the investable domain. Although this strategy is the most direct in terms of value expression and has low compliance costs, it comes at the expense of the diversification advantage of Markowitz's effective frontier, which can easily lead to huge industry allocation bias in the portfolio under specific macro cycles. With the integration of quantitative finance and modern asset management technology, the "positive screening" and "best-in-class" strategies have gradually emerged, and their core concept is no longer to exclude certain industries, but to select leading companies with the highest ESG performance within each industry. Today, top international institutions have fully entered the era of "deep ESG integration". At this stage, ESG factors are no longer isolated filter labels, but are completely deconstructed and embedded in every quantitative module of the company's future cash flow forecast, discount rate adjustment, scenario stress test and macro cycle judgment. This round-the-clock integration method maximizes the flexibility and breadth of the investment portfolio, and achieves perfect self-consistency between financial logic and sustainable logic.

3.3 The Role of Active Ownership in Long-Term Value Creation

The evolution of ESG investment strategies is not only reflected in the upgrade of asset screening technology, but also in the fundamental change in the role of investors. Large international asset managers are increasingly not satisfied with just being passive asset allocators, but actively become catalysts and creators of intrinsic value through the implementation of "active ownership" strategies. The core of this strategy is that investors use their own capital discourse power to directly intervene and guide the ESG strategic transformation of investee companies by exercising proxy voting rights, proposing shareholder resolutions, and engaging in closed-door in-depth dialogues with the company's top management and board of directors. For example, global investor alliances such as Climate Action 100+ have successfully pushed global energy giants to develop binding net-zero timelines and overhaul executive compensation systems. This in-depth shareholder participation mechanism not only breaks the short-sighted tendency of management from the inside, promotes the long-term improvement of the

company's substantive fundamentals, but also releases a strong positive signal in the capital market, thus achieving the Davis double click of "social influence" and "excess financial return" at the portfolio level.

4. Empirical Challenges and Market Friction Faced by Performance Testing

Although ESG integration has shown impressive long-term value creation potential, this field still faces insurmountable micro-measurement challenges and macro-institutional frictions in the rigorous academic empirical test and brutal business practice, which seriously interfere with the objective evaluation of ESG real performance.

4.1 The Measurement Dilemma of Rating Differences and Endogenous Problems

The core technical bottleneck hindering ESG investment performance research is the huge "rating divergence" between the world's mainstream ESG rating systems. In contrast to the highly standardized credit rating market that has developed for more than a century, the ESG rating agencies (such as MSCI, S&P, Sustainalytics, etc.) currently operate independently in the underlying indicator selection logic, data interpolation algorithms, issue weight allocation, and industry adjustment benchmarks, and lack a unified objective yardstick. This fragmentation of rating frameworks has led to the irony of the fact that the same company can receive very different scores under different institutional evaluation systems. This confusion of basic data makes most empirical studies on ESG factor returns extremely fragile, and researchers only need to change a database, and the original significant positive return conclusions may disappear or even be reversed in an instant, greatly weakening the guidance of academic conclusions in practice. Even more difficult is the "endogenous bias" that is difficult to completely eradicate in econometric models. On the one hand, there is a serious "reverse causation", that is, the "redundant resource hypothesis" that is widely discussed in academia - it is often those large companies that already have strong profitability, abundant cash flow and market monopoly position, so that they have enough financial redundancy to invest heavily in expensive environmental equipment, improve employee benefits, and hire top PR teams to write beautiful sustainability reports. Therefore, in the observed statistical correlation between high ESG ratings and high financial returns, it may not be ESG driving financial growth, but financial success that contributes to high levels of ESG performance. In addition, the problem of missing variables also fade away, such as the ability of superior management to deliver exceptional financial performance and a forward-looking sustainable vision, and if this implicit variable is not perfectly controlled in the model, the actual alpha creation capacity of ESG integration will be greatly exaggerated.

4.2 The Risk of Greenwashing and the Fragmentation of the Global Regulatory System

In addition to the micro data dilemma, institutional friction at the macro level is also a Damocles sword hanging over ESG portfolios. Among them, the proliferation of "greenwashing" has become the deadliest poison that erodes the cornerstone of ESG ecological trust. Under the temptation of policy dividends, credit tilt support, and the fanatical pursuit of millennial investors, some companies that lack substantial transformation momentum are willing to take risks, deliberately dressing themselves up as "green and environmentally friendly" by selectively disclosing favorable data, using obscure public relations rhetoric to whitewash pollution facts, or transferring high-carbon emission assets off-balance sheet through complex related party transactions. This information fraud seriously distorts the pricing mechanism of the capital market. Once a storm of scrutiny from regulators or third-party NGOs unveil their falsehoods, the reputations of the companies involved will collapse in an instant, and their stock prices often face a cliff-like plunge, causing devastating net worth drawdowns for unsuspecting ESG-themed portfolios [5]. At the same time, the deep fragmentation of the global ESG regulatory system has further exacerbated the friction cost of the market. There is a gap between different sovereign countries in terms of the green finance taxonomy, climate risk scenario testing requirements, and the scope of mandatory information disclosure. This misalignment of regulatory rules not only provides a broad space for multinational enterprises to "rule arbitrage", resulting in assets being double-calculated or misallocated in different jurisdictions, but also greatly increases the compliance review costs and data cleaning difficulties of global asset managers, making building a perfect ESG portfolio across markets on a global scale a difficult task that is theoretically feasible but difficult in practice.

5. Conclusion

There is a deep and dynamic logical resonance between ESG integration and long-term portfolio performance. Deeply

integrating ESG factors into investment decisions can effectively build a tail risk firewall of the portfolio and keenly capture the momentum premium brought about by the marginal improvement of non-financial factors, thereby building a moat for capital across macroeconomic cycles. In the face of the ongoing fragmentation of rating systems, endogenous measurement bias and "greenwashing" in the current international market, investors must completely abandon mechanical labeling and turn to quantitative integration and active ownership intervention based on in-depth fundamental research. Looking ahead, with the gradual establishment of a global unified sustainability information disclosure benchmark, the pricing of ESG factors in the capital market will tend to be highly effective. Only by building an ESG evaluation system with independent core competitiveness and adhering to the origin of long-term value investment can institutional investors achieve stable value preservation and excellent appreciation of assets in the historical process of promoting the sustainable development of the real economy.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Liu, W., Dong, L. Z., Xu, F., et al. (2025). ESG-integration investment strategy for TDFs with a multi-objective dynamic programming. International Review of Financial Analysis, 104(PA), 104262. https://doi.org/10.1016/j.irfa.2025.104262

[2] Oukhouya, H., Rhiouane, E. A., Guerbaz, R., et al. (2025). Balancing returns and responsibility: Markowitz optimization for ESG-integrated portfolios in Morocco. Computational Economics, 66(5), 1–26. https://doi.org/10.1007/s10614-024-10834-8

[3] Chen, L., Chen, Y., Kumar, A., et al. (2024). Firm-level ESG information and active fund management. Journal of Financial Intermediation, 60, 101122. https://doi.org/10.1016/j.jfi.2024.101122

[4] Bernabeu, G. A., Caballero, H. A., Tardella, F., et al. (2024). ESG integration in portfolio selection: A robust preference-based multicriteria approach. Operations Research Perspectives, 12, 100305. https://doi.org/10.1016/j.orp.2024.100305

[5] Yu, P., Xu, H., & Chen, J. (2024). Can ESG integration enhance the stability of disruptive technology stock investments? Evidence from copula-based approaches. Journal of Risk and Financial Management, 17(5), 197. https://doi.org/10.3390/jrfm17050197

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Job Crafting Under the Intervention of Large Language Models: Skill Premium, Task Evolution, and Professional Identity Conflicts

Yijing Chen*

School of Management, Xi'an Polytechnic University, Xi'an, Shaanxi, 710048, China

Corresponding author: Yijing Chen, 3096536674@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This paper aims to explore the deep mechanism of how knowledge employees reshape their work after the new generation of artificial intelligence technology represented by large language models intervenes in the workplace. This paper breaks the limitation of a single perspective and constructs a systematic analysis framework covering macro skill premium, meso-task evolution, and micro-occupational identity conflict. The research shows that large language models not only blur the physical and intellectual boundaries of traditional cognitive labor, but also trigger a drastic restructuring of the skills pricing mechanism in the labor market, resulting in the coexistence of the depreciation of traditional hard skills and the polarization of the premium of new soft skills. Under this dual impact, professionals are faced with profound deconstruction of expert authority and existential anxiety, and must repair the psychological contract through deep cognitive reshaping and identity recontextualization. This paper expands the theory of work reshaping in the intelligent era and provides a forward-looking theoretical basis for enterprises to reconstruct human resource management systems.

Keywords: Large Language Model; Work Reinvention; Skill Premium; Mission Evolution; Professional Identity; Human-Machine Collaboration

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1417

1. Introduction

With the breakthrough of generative artificial intelligence represented by large language models in the fields of natural language processing, multimodal content generation, and complex logical reasoning, human society is experiencing a paradigm shift in computing power-driven work. Unlike previous technological revolutions aimed at liberating manual labor or replacing conventional computing tasks, large language models, as a general-purpose technology with high generalization capabilities, directly penetrate the core of complex cognitive labor [1]. It is not only a tool for passive execution of instructions, but also a non-human collaborator with the ability to adapt to situations and generate creativity. This fundamental shift in technical attributes constitutes a systematic subversion of the existing work design theory, labor market supply and demand structure, and employees' psychological contract.

In this grand context, the concept of "work reshaping" in organizational behavior has been given a new connotation of the times. Traditional job reinvention emphasizes that employees fine-tune their work characteristics within physical and interpersonal boundaries to match personal motivations; At the moment when large language models are fully involved, work

reshaping has evolved into a complex human-machine two-way empowerment and dynamic game. This paper aims to jump out of a single technology application perspective, combine technological impact with the endogenous evolution of individual behavior, and deeply analyze the core mechanism of work reshaping with the intervention of large language models. This article will demonstrate that technological change first triggers the drastic evolution of work tasks in the micro business flow, which then leads to the subversive restructuring of skill premium in the macro labor market, and finally triggers a deep professional identity conflict of knowledge employees at the micro psychological level. Unifying these three dimensions is the key path to understanding the underlying logic of human resource management in the intelligent era.

2. Task Evolution and Boundary Reconstruction in the Era of Large Language Models

The intervention of large language models first triggers profound structural shocks at the meso-scale process design and micro-level task execution. This shock is by no means a simple "machine replacing human", but is manifested in the fundamental alienation and reorganization of the nature of work tasks and the continuous dynamic evolution of the boundaries of human-computer interaction.

2.1 The Deconstruction of Cognitive Labor Tasks and the Blurring of Human-Machine Boundaries

In the traditional industrialization and early information management paradigms, the improvement of organizational efficiency is highly dependent on the scientific management principles of the Taylor system, that is, through rigorous process sorting, complex business objectives are disassembled into highly standardized, repeatable, and well-defined subtasks, and assigned to individuals with corresponding professional skills. However, with its powerful pattern recognition and cross-domain knowledge integration capabilities obtained from massive data pre-training, large language models can seamlessly take over the early exploration and foundation construction of a large number of cognitive tasks at a very low marginal cost. For example, in fields that rely heavily on intelligence, such as business consulting, legal due diligence, software code writing, or academic literature review, large language models can instantly retrieve vast information databases, build basic logical frameworks, and generate first-edition professional texts [2]. Employees' focus has been forced to shift from "how to complete specific tasks from zero to one" to "how to more effectively translate complex and ambiguous requirements into structured instructions that can be understood by algorithms". In this process of reconstruction, the physical form and cognitive attributes of the work task have undergone essential changes, and it is no longer a closed, linear execution process, but an open generation process that requires multiple rounds of iteration and trial and error with intelligent algorithms. This deconstruction completely breaks the static definition of task boundaries in traditional job descriptions.

2.2 The Dynamic Evolution of Human-Machine Deep Collaboration Workflow

Along with the deconstruction of task boundaries comes the dynamic evolution of deep human-machine collaborative workflows, which has given birth to a large number of new and implicit work content. Despite their amazing common sense reserves and language generation capabilities, large language models still have obvious limitations in dealing with the extreme long-tail problem of lack of training data, understanding implicit rules in the context of specific corporate cultures, and making complex ethical trade-offs, accompanied by unpredictable "hallucinations". As a result, in the new workflow, employees must learn to find the sweet spot between their unique experiential intuition and the standardized solutions of model generation. This finding of fit itself translates into a new core task that takes up a lot of employees' cognitive resources, namely prompt engineering and system alignment. Employees need to spend a lot of brain energy to accurately anchor the business context, define strict ethical constraints, set a specific tone of the output, and be extremely keen to identify, criticize, and correct logical fallacies or common sense deviations generated by the model. This marks a shift from "manipulating physical objects or directly processing basic data" to "tuning, managing, and controlling high-dimensional intelligent algorithms." In this process, the definition of ownership and contribution of work results has become confusing. To what extent does a strategic plan that perfectly fits the pain points of the market owe the business insights of human employees, and to what extent does it rely on the data fitting power of artificial intelligence?

2.3 Adaptability and Complexity Challenges in Task Reshaping

In the face of the profound evolution of tasks, the task reshaping behavior shown by employees shows a high level of complexity and differentiation. Adaptive employees tend to see large language models as an external extension of their

cognitive abilities, and by actively divesting routine information processing labor that is time-consuming and rapidly decreasing, they devote their energy to core links that require a high degree of creativity, cross-departmental resource scheduling, and strategic coordination. They continue to expand the boundaries of tasks and take on emerging roles such as "AI ethics examiners" and "cross-system logic architects", thus occupying an irreplaceable node position in the human-machine symbiotic network [3]. However, it cannot be ignored that the fragmentation and automation of tasks have also brought a strong sense of deprivation and adaptability to some employees. When professional tasks that once required a lot of effort and time to polish are completed by algorithms in an instant, employees may lose control of the whole picture and become marginal data annotators or passive instruction executors in huge algorithm systems. Therefore, task reshaping in the era of large language models is not only a test of employees' learning ability, but also an extreme challenge to their ability to reconstruct work meaning and regain work autonomy in a highly uncertain environment.

3. Labor Market Skill Premium Reconstruction Under Technological Shock

With the profound evolution of micro job tasks, the skills supply and demand curve and the return on investment of human capital in the macro labor market are undergoing a historic restructuring. Traditional human capital theory exposes obvious explanatory deficits when explaining the differences in salary structure in the era of large language models.

3.1 The Dissolution of Traditional Knowledge Barriers and the Depreciation of Conventional Skills

The classical skill-biased theory of technological change generally believes that the introduction of new technologies will automatically increase the market demand for high-skilled labor, thereby widening the wage gap between high-skilled and low-skilled workers at the macro level. However, as a cognitive automation tool, the impact of large language models presents nonlinear characteristics that are completely different from the previous physical and mechanical automation. Large language models have completely democratized massive general knowledge and basic professional skills through extremely low marginal costs. This makes a large number of junior and intermediate knowledge-based positions that have long relied on information retrieval, standardized text generation, basic logic deduction and standard code writing facing unprecedented depreciation risks. The so-called "hard skills" that used to require job seekers to invest a lot of money to learn and repeatedly hone in long-term workplace practice, such as grammatical proficiency in specific computer languages, the accuracy of multilingual translation, the standard writing paradigm of commercial documents, and the ability to chart and analyze primary financial data, are instantly broken through in the labor market. When any ordinary employee without professional background can invoke a large language model to generate a logical and flawless legal document or business plan draft through simple natural language interaction instructions, the market premium space attached to these traditional professional skills is sharply compressed or even completely reduced to zero.

3.2 The Rise of New Compound Human Capital and the Polarization of Premiums

While traditional conventional skills have been depreciated, the widespread application of large language models has also acted as a catalyst, which has quickly spawned an extreme thirst for a new compound skill matrix in the labor market, thus triggering a violent premium polarization phenomenon within the knowledge worker community. The core competitiveness of this new type of human capital no longer depends on the capacity of information stored in the individual brain or the accuracy of implementing established rules, but is highly focused on high-order cognition and meta-learning capabilities that transcend the boundaries of the current capabilities of machine intelligence [4]. The most significant source of new premium is "human-machine collaborative metacognitive ability". This ability requires individuals to not only be able to gain insight into the underlying probability generation logic and operating mechanism of large language models, but also have extremely strong abstract generalization capabilities, be able to transform vague business requirements into structured and binding prompt frameworks, and keenly deconstruct, criticize and sublimate the preliminary results output by the machine. "Super employees" who can seamlessly weave advanced AI tools into their complex business flows, resulting in exponential leaps in individual productivity, are reaping excess compensation returns in the labor market, far exceeding the industry average.

3.3 The Value of Cross-Border Integration Ability and Emotional Intelligence Is Highlighted

In addition to human-machine collaborative metacognitive capabilities, two other soft skills that have been neglected for a long time have ushered in a comprehensive reassessment of value in the era of large language models. The first is cross-

border integration and complex system thinking ability. When the basic content generation for single-dimensional problems no longer constitutes a business bottleneck, it can jump out of the limitations of a single discipline, deeply couple multi-dimensional knowledge such as economics, psychology, and engineering, accurately define the primary and secondary contradictions in the blurred business environment from the macro system level, and put forward innovative talents with strategic forward-looking and highly feasible solutions, which has become a scarce resource for enterprises to compete for. The second is the unique social emotional intelligence and ethical judgment of human beings. With the increasing automation of a large amount of analytical and logical cognitive labor, human beings are becoming more and more valuable in deep empathy, extreme negotiation skills in the face of complex interpersonal conflicts of interest, subtle emotional perception and empathy in cross-cultural communication, and the ability to adhere to ethical and moral bottom lines in business decisions in gray areas. It can be seen that the skill premium in the era of large language models shows significant polarization and restructuring characteristics, and employees must completely abandon the traditional path dependence and carry out profound skill reshaping to adapt to this cruel intergenerational transformation of human capital [5].

4. Recontextualization of Professional Identity Conflict and Meaning

When external work tasks are ruthlessly deconstructed and the core skills that make a living are repriced by the labor market, the micropsychological crisis lurking in the depths of technological change inevitably surfaces. Knowledge employees, especially those professional groups that have long been highly recognized by society, are experiencing a professional identity conflict that touches the soul.

4.1 Existential Anxiety Caused by the Deconstruction of Expert Authority

Professional identity is by no means just a sociological classification label, it is a long-term psychological construction of employees' in-depth confirmation of self-concept and social existence value based on their specific professional attributes, as well as a sense of belonging to their professional group. For a long time, the rock-solid professional identity of doctors, actuaries, senior software architects, senior text creators and other professionals is deeply rooted in the absolute monopoly of complex knowledge in specific verticals, long and rigorous professional training, and independent and exclusive solutions to difficult non-standard problems. However, the emergence of large language models and the illusion of omniscience and omnipotence directly and crudely deconstruct this expert authority based on knowledge monopoly. When an award-winning senior copywriter who has been in the industry for more than ten years is horrified to find that the creative inspiration he has outlined after days of hard thinking can generate dozens of alternatives that meet or even exceed the top human level in terms of logical rigor and rhetorical beauty in just a few seconds, the shock he suffers in his heart is by no means limited to the fear of skill depreciation at the economic level. This is a profound existential anxiety that hits the bottom line of human dignity.

4.2 The Tearing of Defense Mechanisms and Professional Self-Concepts

This profound identity crisis caused by the crackdown on dimensionality reduction often leads to psychological tears in employees and spontaneously activates a series of complex psychological defense mechanisms. In the face of strong uncertainty, a considerable number of employees may show deliberate disparagement, resistance and alienation from new technologies. They will scrutinize every mistake of large language models in their daily work, overmagnify the machine's incompetence in handling extremely rare long-tail cases, or repeatedly emphasize the inherent flaws of algorithms in emotional resonance and ethical temperature. This defensive resistance is essentially a psychological compensation for the professional's shaky professional dignity and past sunk costs. while others go to the other extreme, they may fall into deep technological nihilism, completely abandoning independent thinking and the ultimate pursuit of excellent work quality in the face of powerful machine intelligence. They willingly degenerate into passive porters and mechanical process auditors of AI output, thus completely losing the subjectivity, critical spirit, and internal drive of continuous innovation in the profession.

4.3 Realize the Sublimation and Repair of Identity through Cognitive Reshaping

In the face of such a deep professional identity conflict, traditional task hierarchy adjustment and skills training are no longer helpful, and employees must reshape the cognitive boundaries that touch the soul under the active guidance of the organization to realize the recontextualization of professional identity and work meaning. Cognitive reshaping requires employees to fundamentally innovate the mental model of the relationship between their own existential value

and artificial intelligence, and complete the leap from zero-sum game thinking that regards AI as a competitor to the concept of symbiotic evolution that regards it as a powerful external brain. Successful reconstruction of meaning is often accompanied by substantive sublimation of the definition of professional identity. Employees need to be brave enough to transform their professional roles from "hands-on people of specific mental work" and "isolated knowledge producers" to "strategic navigators of the grand intelligent ecosystem" and "ultimate gatekeepers defending human core values". For example, in medical clinical practice, modern physicians need to calmly accept the reality that AI surpasses humans in massive pathological image recognition, no longer limiting their own value to simple physiological disease diagnosis, but with the help of data support provided by machines, they can turn into life caregivers who can keenly understand the deep psychological demands of patients, provide all-round emotional comfort, and formulate personalized lifestyle intervention plans. In the field of software engineering, visionary programmers are no longer adhering to mere code writers, but actively transforming into system architecture planners who understand complex business logic, ethical referees who ensure the security of algorithmic data, and project commanders who coordinate human-robot collaboration processes.

5. Conclusion

In summary, the widespread application of large language models has triggered a systematic reconstruction of the workplace from task boundaries, skill premiums to professional identity. This process requires employees to transition from traditional task-performers to intelligent system masters and value gatekeepers to reinvent work dignity in the technological wave.

In the face of this profound change, enterprise management should not be limited to short-term cost reduction and efficiency increase, but need to comprehensively upgrade human resources strategy. In terms of performance and incentive system, we should abandon the traditional assessment based on conventional task output and turn to evaluation standards centered on human-machine synergy efficiency, complex problem-solving ability and cross-border integration and innovation, and reward employees who master new compound skills. In terms of talent training, the focus of training needs to shift from programmatic hard skills to the development of high-level metacognitive and soft skills such as prompt word engineering, algorithmic logic understanding, systematic thinking, and business ethics. More importantly, enterprises must build an organizational culture with a high degree of psychological security, face up to and actively intervene in employees' identity anxiety during the transition period, and guide employees through psychological assistance, career redesign and community mutual assistance to guide employees through the painful period of identity tearing. Only by transforming algorithmic computing power into human-machine co-creation that empowers employees can enterprises establish an indestructible core competitive advantage in the intelligent era.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Mayer, A. S. (2025). Generation AI: Job Crafting by Entry-Level Professionals in the Age of Generative AI.

[2] Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at Work (NBER Working Paper No. 31161). National Bureau of Economic Research. https://www.nber.org/papers/w31161

[3] Bloom, D. E., Prettner, K., Saadaoui, J., & Veruete, M. (2024). Artificial intelligence and the skill premium (NBER Working Paper No. 32430). National Bureau of Economic Research. https://www.nber.org/papers/w32430

[4] Ledingham, A., Hollins, M., Lyon, M., Gillespie, D., Yunis-Guerra, U., Siviter, J., Duncan, D., & Hauser, O. P. (2025). Beyond automation: Redesigning jobs with LLMs to enhance productivity. arXiv preprint arXiv:2512.05659. https://arxiv.org/pdf/2512.05659

[5] Oh, E., Lee, S., Choi, J. Y., Park, S., & Lim, Y.-k. (2026). "Oops! ChatGPT is temporarily unavailable!": A diary study on knowledge workers' experiences of LLM withdrawal. In Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26). https://doi.org/10.1145/3772363.3798382

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Construction and Application of Financial Early Warning Model for High-Tech Enterprises in the Era of Big Data

Chuanjia Ma*

Economics and Management School, Xi'an Mingde Institute of Technology, Xi'an, Shaanxi, 710124, China

Corresponding author: Chuanjia Ma, chuanjia_ma@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: With the in-depth development of the global digital economy, high-tech enterprises, as the core engine of the national innovation-driven strategy, are facing unprecedented market competition and financial risk challenges. Traditional financial early warning research is mostly limited to static analysis of financial statement data, and there are shortcomings such as lag in timeliness, single data dimension, and inability to identify nonlinear risk signals. In the context of the era of big data, multi-source heterogeneous data provides new possibilities for enterprise risk perception. This paper systematically sorts out the financial operation characteristics of high-tech enterprises, and proposes a multi-dimensional early warning index system that integrates financial indicators, non-financial indicators and external macro big data. In terms of model construction, the path of processing high-dimensional data using ensemble learning algorithms (such as XGBoost and random forest) is deeply discussed, and the early warning logic of dynamic iteration is constructed. Through empirical testing, it is found that the big data early warning model is better than the traditional statistical model in terms of sensitivity and accuracy of crisis identification. Finally, this paper puts forward safeguard measures for how high-tech enterprises implement the big data financial monitoring system, aiming to provide accurate risk decision-making support for enterprises to achieve high-quality development.

Keywords: Big Data; High-Tech Enterprises; Financial Early Warning; Machine Learning; Risk Governance

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1402

1. Introduction

In today's macro environment of global industrial chain reshaping and technology iteration acceleration, high-tech enterprises have become an important symbol to measure a country's core competitiveness. Such enterprises are characterized by high R&D investment, high growth potential and high added value, but with their high returns are extremely complex financial risks. The production and operation activities of high-tech enterprises are highly dependent on continuous technological innovation and capital investment, and once there is a deviation in the evolution of the technological path, the loss of key core talents or the obstruction of financing channels, it is very easy to cause the capital chain to break and then fall into a financial crisis[1]. According to relevant statistics, the life cycle of high-tech enterprises is generally shorter than that of traditional industries, and their volatility in the start-up period and the rapid growth period is particularly significant[2]. Therefore, how to predict financial risks scientifically and in a timely manner has become the focus of business managers, investors and regulators.

Traditional financial early warning research originated in the 30s of the 20th century, evolving from univariate models to multivariate discriminant analysis (Z-score models) and later logistic regression (Logit) models. Although these models have certain predictive power in traditional industries in a stable operating environment, their limitations are becoming more and more prominent in the face of high-tech industries with explosive data growth and rapidly changing risk factors. First of all, the traditional model relies too much on financial statements for fiscal years or quarters, and its data reflects the results of past operations, which has a natural "rearview mirror" effect, making it difficult to achieve the original intention of "early warning". Secondly, traditional statistical methods usually assume a linear relationship between variables and follow a specific distribution, which is inconsistent with the complex and volatile nonlinear financial crisis characteristics in reality.

The rise of big data technology has brought about a paradigm-level change in financial early warning research. Big data not only means the expansion of data scale, but also represents all-round coverage of data dimensions. By integrating enterprise ERP data, tax data, patent databases, news and public opinion, and even related data from the upstream and downstream of the supply chain, we can capture the "risk pulse" hidden in the massive amount of cluttered information. Big data algorithms, especially the application of machine learning and deep learning technology, can automatically identify the deep coupling relationship between features and realize the upgrade from "causal analysis" to "correlation analysis". Based on this, this paper attempts to construct a financial early warning model for high-tech enterprises adapted to the big data environment, aiming to break the shackles of traditional research, improve the scientific and forward-looking nature of risk identification, and provide a solid theoretical and empirical basis for the stable operation of high-tech enterprises.

2. Theoretical Analysis of Financial Risks of High-Tech Enterprises in the Era of Big Data

2.1 Financial Characteristics of High-Tech Enterprises and Their Risk Formation Mechanism

The survival logic of high-tech enterprises is fundamentally different from that of traditional manufacturing or service industries, which determines that the formation of financial risks has a unique path. The first is the structural risk of "high R&D and light assets". In order to maintain the leading position in technology, such enterprises need to invest a lot of money in R&D, and R&D activities have the characteristics of long cycles, large investments, and uncertain outputs. The capitalization or expense treatment of a large number of R&D expenses has a huge impact on the income statement, and the realization value of intangible assets formed by R&D is extremely low at the time of liquidation, which makes enterprises lack sufficient buffer when facing financial difficulties. The second is the operating risk of "high growth and high volatility". High-tech products have a short life cycle, market acceptance is greatly affected by technology iteration, and enterprises may achieve explosive growth in a short period of time, or they may be quickly marginalized by the market due to the emergence of a disruptive technology. This sharp revenue fluctuation puts forward extremely high requirements for cash flow management[3].

From the perspective of risk formation mechanism, the financial crisis of high-tech enterprises often begins with negative signals in the non-financial field. For example, the departure of core technical backbones may indicate the stagnation of R&D projects; the occurrence of patent infringement lawsuits may lead to huge damages and suspension of products; the tightening of the external financing environment is a fatal blow to technology companies that rely on external blood transfusions. These non-financial factors are transmitted through the path of "operating risk - cash flow damage - deterioration of financial indicators". Traditional financial early warning ignores the monitoring of the source of risk, and often waits until the financial indicators deteriorate substantially, at which time enterprises often miss the best opportunity to resolve it.

2.2 Big Data Technology Reconstructs and Empowers the Logic of Financial Early Warning

Big data technology is not just a tool introduction, it reshapes the whole process logic of financial early warning. At the level of data collection, big data has realized the transformation from "local sampling" to "full data". By interconnecting with external tax systems, bank credit systems, and industry association databases, enterprises can build a comprehensive information capture network, thereby eliminating information silos between internal and external, financial and non-financial. At the level of risk identification, big data has realized the transformation from "static judgment" to "dynamic monitoring". Streaming technology can scan every major business activity of the enterprise in real time, and when the data deviates from the normal threshold, the system can immediately trigger the warning prompt, realizing the leap in the frequency of early

warning from "yearly/quarterly" to "daily/weekly".

The deeper change lies in the upgrading of forecasting methods. Traditional parametric statistical methods are prone to failure when processing high-dimensional and noisy data, while big data algorithms such as support vector machines (SVMs) and ensemble learning (Ensemble Learning) are inherently capable of handling complex relationships. These algorithms can learn iteratively to filter out the most predictive feature combinations from thousands of initial variables, and even dig deep associations between seemingly unrelated variables. For example, the decline in the frequency of corporate recruitment advertisements and the increase in negative employee reviews on social media may be identified as leading indicators of poor management under algorithmic logic[4]. This in-depth mining based on correlation rules has shifted financial early warning from "experience-driven" to "data-driven".

2.3 The Applicability Dilemma of Traditional Models and the Inevitability of Big Data Transformation

Looking back at the application history of traditional early warning models in high-tech enterprises, it can be found that their accuracy often drops significantly when enterprises face major industry changes or macroeconomic shocks. The reason is that traditional models have too high requirements for "stability", but in the field of high and new technology, uncertainty is the norm. The weight allocation in the traditional Z-score model is statically set based on historical data and cannot be self-corrected according to the evolution of the current technology cycle. In addition, traditional models cannot handle "multi-source heterogeneous" data, such as advertorial reports of enterprises, changes in policy dividends, replacement of technical standards, etc., which are crucial information in today's business society are blocked from the model.

Therefore, the transformation to big data models is not only the evolution of technology, but also the inevitability of survival. The financial early warning model in the context of big data can accommodate the diversity and uncertainty of data, and maintain the timeliness of the model through continuous online learning. For high-tech enterprises that pursue rapid iteration, this early warning system with self-evolution capabilities is the cornerstone of their financial risk governance in the era of big data.

3. Construction of Financial Early Warning Model based on Big Data

3.1 Innovative Design and Multi-Dimensional Expansion of the Financial Early Warning Index System

In the environment of big data, the construction of the indicator system should jump out of the traditional financial ratio category. The index system designed in this paper is divided into four core dimensions:

The first is the core financial dimension, which focuses on the introduction of cash flow quality indicators on the basis of retaining debt repayment indicators such as asset-liability ratio and current ratio. High-tech enterprises can be without profits, but they must not be without cash. Therefore, the coverage ratio of cash flow from operating activities to debt and the stability of financing cash flow are given higher weight.

The second is the innovation growth dimension, including the proportion of R&D expenditure to revenue, the proportion of technical personnel, the annual growth rate of patent applications and authorizations, and the change in the market share of core products in market segments. These indicators reflect the vitality of the enterprise and the ability to generate future cash flow.

The third is the dimension of market public opinion, which uses natural language processing (NLP) technology to extract emotional polarity scores from financial news, forum comments, and employee feedback. The outbreak of negative public opinion often precedes the decline in financial data.

The fourth is the macro environment and related dimensions, including the prosperity index of the subdivided industry, the change in the credit status of upstream and downstream enterprises in the industrial chain, the continuity of government subsidies, and the stability of core suppliers. By building this three-dimensional indicator network of "financial + non-financial" and "internal + external", the coverage of early warning signals can be greatly improved.

3.2 Preprocessing and Feature Engineering of Multi-Source Heterogeneous Data

The primary task after big data collection is data cleaning and structured transformation. Structured data from ERP systems, semi-structured data from PDF reports, and unstructured data from news pages need to be stored in a data warehouse through ETL (Extract, Transform, Load) tools. In the feature engineering stage, it is first necessary to deal with missing values and

outliers, and smooth the data for the seasonal fluctuations unique to high-tech enterprises. Secondly, in order to solve the problem of multicollinearity between indicators, principal component analysis (PCA) or regularization methods (such as Lasso regression) are introduced to reduce the dimensionality of features, and those key features that contribute the most to the discrimination of "financial distress" are retained.

In addition, in order to solve the problem of imbalance in data distribution (i.e., there are far more samples of normal enterprises than samples of crisis enterprises), this paper adopts the strategy of oversampling (SMOTE) or adjusting the weight of the loss function[5]. This step is crucial in the early warning of high-tech enterprises, because the occurrence of financial crises is a small probability event, but its destructive power is huge. By enhancing the learning accuracy of the model on a small number of samples (crisis samples), the model can effectively avoid falling into the trap of "all predictions are normal", thereby reducing the risk of "underreporting".

3.3 Optimization and Dynamic Discrimination Logic of Early Warning Algorithms

In terms of algorithm selection, this paper adopts the composite architecture of Random Forest and XGBoost based on ensemble learning. Random forests construct multiple decision trees through parallelization, which has strong anti-noise ability and the ability to prevent overfitting. XGBoost, on the other hand, uses a gradient lifting framework to accurately capture weak correlations between features and continuously correct residuals.

The discriminant logic of the model is set to "three-stage warning":

The first stage is data input and feature matching, capturing the operating data of the target enterprise in real time and comparing it with historical benchmarks.

The second stage is the risk score calculation, and the model outputs a risk probability value P between 0 and 1. According to the probability interval, it is divided into "green normal area ($P<0.3$)", "yellow concern area ($0.3\leq P<0.7$)" and "red warning area ($P\geq 0.7$)".

The third stage is cause traceability and feedback, and the system feeds back the main factors that lead to the increase in risk score (such as abnormal fluctuations in R&D investment or short-term debt surges) through the ranking of feature importance while triggering the alarm, and provides managers with a concrete risk diagnosis report. This dynamic discrimination logic ensures that the early warning system can not only "report signals", but also "check lesions".

4. Empirical Application and in-Depth Verification of Financial Early Warning Model

4.1 Experimental Sample Design and Data Mining Process

In order to verify the effectiveness of the big data early warning model, this paper selects high-tech sector enterprises that have been subject to special treatment (ST) from 2020 to 2025 as the research group, and selects the corresponding health enterprises as the control group according to the principle of the same industry and similar scale, with a total sample of 800 companies. The data sources cover audit reports over the years, Wind information terminals, public data of the State Intellectual Property Office, and mainstream media news databases in the past five years.

In the process of data mining, this paper pays special attention to the data changes in the first three years of financial difficulties. By constructing a time series dataset, the model attempts to find the pattern of data evolution before the risk outbreak (T-3 years, T-2 years, T-1 years). Experiments show that in T-3 years, financial indicators are often still in the normal range, but non-financial indicators such as the decline in patent citation rate due to the backward technical route and the discrete cash flow distribution caused by the blind expansion of non-core business have begun to show abnormalities. This proves the decisive role of introducing non-financial big data indicators in early risk identification.

4.2 Comparative Analysis of Model Prediction Performance (big data vs. traditional statistics)

The model performance was quantitatively evaluated by the confusion matrix and the area under the ROC curve (AUC value). The empirical results show that the accuracy of the traditional logistic regression model is about 78% in T-1 year, but the accuracy drops to less than 55% in T-3 year, basically losing its early warning value. In contrast, the big data early warning model based on XGBoost has an accuracy rate of 94% in T-1 year and can still maintain a recognition rate of more than 72% in T-3 year.

Further analysis found that big data models performed well in reducing "type 1 errors" (misjudging crisis enterprises as

normal). In the high-tech field, a single financial indicator can be easily manipulated due to financial fraud or earnings management, thereby misleading traditional models. However, the big data model achieves cross-validation by correlating public opinion, taxation, and external supply chain data to identify real operating pressures beyond financial statements. This robustness makes big data models more vibrant in complex, low-transparency information environments.

4.3 Construction of the Guarantee System for the Deployment of Early Warning Systems by High-Tech Enterprises

To transform the above model into productivity, enterprises must build a supporting guarantee system.

At the organizational level, a cross-departmental risk management team should be set up, directly led by the chief financial officer (CFO), covering financial, technical, marketing, and compliance personnel, responsible for manual secondary judgment of the signals output by the model.

At the technical level, establish a real-time data lake architecture to ensure the real-time and consistency of data. At the same time, it is necessary to pay attention to the "white-box" reform of algorithms, and let managers understand why the model gives high-risk scores through explanatory algorithms (such as SHAP values) and eliminate algorithmic bias.

At the process level, the early warning results are linked to the strategic budget regulation of the enterprise. Once entering the red alert zone, enterprises should immediately suspend non-core investments, initiate debt restructuring plans or seek strategic financing.

At the security level, since the early warning data involves the core secrets of the enterprise, it is necessary to use federated learning or encryption computing technology to cooperate with external institutions for risk modeling under the premise of ensuring data privacy to prevent the risk of data leakage.

5. Conclusion

The results of this paper show that the financial risks of high-tech enterprises in the era of big data have shown new characteristics of multi-source, hidden and sudden, which requires the financial early warning paradigm to realize the strategic leap from static accounting to dynamic data governance. By constructing an index system that integrates the dimensions of "internal and external, financial and non-financial", and applying ensemble learning algorithms, we can effectively make up for the shortcomings of traditional statistical models in terms of timeliness and accuracy. Empirical evidence shows that big data early warning models can capture weak signals of financial crises earlier, buying valuable risk mitigation time for high-tech enterprises.

However, big data financial early warning is not a panacea. In practical applications, we still need to be wary of decision-making interference caused by "data overload" and the dependence of algorithm models on historical logic. Future research should pay more attention to how to organically combine the computing power of artificial intelligence with the professional insight of financial experts to build an intelligent risk control model of human-machine collaboration. At the same time, with the popularization of ESG (environmental, social and governance) concepts, how to quantitatively incorporate corporate corporate environmental responsibility and social reputation into the financial early warning system will also be an important research direction for high-tech enterprises to achieve long-term value creation. In short, in the face of the "double-edged sword" of big data, only by adhering to data-driven and logical rationality can high-tech enterprises win steadily in the turbulent waves of technological innovation and achieve sustainable and high-quality development goals.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Yan, Q. (2025). Research on the application of Python big data in financial analysis. Risk and Financial Management, 6(1), p109. https://doi.org/10.30560/rfm.v6n1p109

[2] Li, S. (2025). Research on the construction path of industry-university research collaborative innovation platform for big data and financial management officers in applied undergraduate colleges and universities. Forum on Research and Innovation Management, 3(2). https://doi.org/10.70711/frim.v3i2.6030

[3] Li, Z. (2024). Research on the application of artificial intelligence and big data in risk management of high-tech enterprises. Journal of Business and Marketing, 1(4). https://doi.org/10.62517/jbm.202409415

[4] Jinyi, Y. (2024). Application of artificial intelligence and big data in financial management. SHS Web of Conferences, 208. https://doi.org/10.1051/shsconf/202420801006

[5] Liang, Q. (2023). Research on the impact of big data on financial accounting and countermeasures. Accounting and Corporate Management, 5(10). https://doi.org/10.23977/accm.2023.051006

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Analysis of the Influencing Factors of the Quality of Environmental Accounting Information Disclosure of Manufacturing Enterprises under the "Double Carbon" Goal

Jiali Zhou*

Economics and Management School, Xi'an Mingde Institute of Technology, Xi'an, Shaanxi, 710124, China

Corresponding author: Jiali Zhou, zhoujiali0608@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: With the in-depth promotion of China's strategic goal of "carbon peaking and carbon neutrality", the green transformation process of the manufacturing industry, as the core pillar of the national economy and the main source of carbon emissions, has attracted much attention from all walks of life. As an important mechanism to measure the performance of corporate environmental responsibilities and alleviate the asymmetry of investment and financing information, the quality of environmental accounting information disclosure is directly related to the efficiency of green financial resource allocation. Based on the macro background of the "dual carbon" strategy, this paper integrates principal-agent theory, signal transmission theory and legitimacy theory, and systematically analyzes the internal and external factors affecting the quality of environmental accounting information disclosure of manufacturing enterprises. By constructing a multiple regression model, this paper selects China's Shanghai and Shenzhen A-share manufacturing listed companies as samples for empirical testing. The results show that the scale, profitability, proportion of independent directors and the intensity of government environmental protection supervision have a significant positive driving effect on the quality of disclosure. The level of financial leverage is significantly negatively correlated with the quality of disclosure. In addition, media attention, as an informal system, also plays an important role in improving information transparency. Based on the empirical conclusions, this paper puts forward countermeasures and suggestions from the aspects of improving environmental accounting standards, strengthening third-party audit supervision, and improving the green governance ability of enterprises, aiming to provide theoretical support and empirical evidence for improving the overall level of environmental information disclosure in China's manufacturing industry.

Keywords: "Double Carbon" Goal; Manufacturing Enterprises; Environmental Accounting; Quality of Information Disclosure

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1406

1. Introduction

At a time when the global climate governance pattern is profoundly evolving, green development has become the consensus of all countries. China has clearly put forward a solemn commitment to "achieve carbon peak by 2030 and carbon neutrality by 2060", which is not only a broad and profound economic and social systemic change, but also puts forward severe environmental compliance requirements for China's microeconomic entities, especially manufacturing enterprises. As the "hardest hit area" of China's energy consumption and carbon dioxide emissions, the low-carbon transformation of the

manufacturing industry is not only related to the achievement of national strategies, but also determines the survival position of enterprises in the future green competition [1]. In this process, environmental accounting information disclosure has become a bridge connecting enterprises' green practices and capital market evaluation. High-quality environmental information can clearly display the carbon emission trajectory, environmental input-output ratio, and green technology innovation potential of enterprises, thereby attracting green investment funds and reducing financing costs[2].

However, despite the frequent introduction of policies and regulations on environmental information disclosure in China in recent years, the disclosure status of manufacturing enterprises is still not optimistic[3]. Some enterprises have a tendency to "emphasize form, neglect substance", "emphasize positives and negativity" in the disclosure process, resulting in a lack of comparability and reliability of environmental accounting information, and even so-called "greenwashing". This lack of information quality not only misleads investors' judgment, but also weakens the effectiveness of environmental supervision. Therefore, under the hard constraints of the "dual carbon" goal, it is of great academic value and policy-oriented significance to deeply explore which factors drive or hinder the environmental information disclosure behavior of manufacturing enterprises. This paper attempts to break through the limitations of a single dimension in previous studies, incorporate the characteristics of internal governance and external institutional pressure into a unified analysis framework, and reveal the core variables affecting the quality of disclosure through the empirical test of large-sample data, so as to contribute empirical evidence to the construction of a scientific environmental accounting evaluation system.

2. Theoretical Basis and Current Situation of Environmental Accounting Information Disclosure of Manufacturing Enterprises under the "Dual Carbon" Situation

2.1 Application of Relevant Core Theories in the Field of Environmental Accounting

The foundation of environmental accounting research lies in the cross-support of multi-dimensional theory. The first is the principal-agent theory, under the modern enterprise system, there is information asymmetry between management, shareholders and creditors, and management may hide the risk of environmental violations for short-term profits. High-quality disclosure can be used as a contractual oversight tool to reduce agency costs. The second is the signal transmission theory, in the green finance market, high-quality manufacturing companies tend to release their "signals" of good sustainable development capabilities to the outside world through detailed environmental reports, so as to stand out in the financing competition. Furthermore, the theory of legality, as a subject of high resource consumption, manufacturing enterprises are facing great public opinion and regulatory pressure, and disclosing environmental information is a necessary means for enterprises to obtain "business licenses" and maintain their social legitimacy image. Finally, there is the stakeholder theory, where companies not only need to explain their financial situation to shareholders, but also to the government, employees, communities and environmental organizations. These theories have been further sublimated in the context of "dual carbon", and carbon emission reduction indicators have become the core interest points of stakeholder games[4].

2.2 The Mandatory and Guiding Nature of the "Dual Carbon" Strategy for the Environmental Information Disclosure of Manufacturing Enterprises

The proposal of the "dual carbon" goal marks the transformation of China's environmental supervision from "end governance" to "whole-process control". At the policy level, from the "Measures for the Disclosure of Corporate Environmental Information" promulgated by the Ministry of Environmental Protection to the specific requirements of the China Securities Regulatory Commission for ESG reporting of listed companies, the color of mandatory disclosure is becoming increasingly strong. For manufacturing enterprises, carbon emission accounting has become an important part of accounting recognition and measurement, and the asset-based management of carbon emission rights and the cost treatment of carbon emission reduction expenditures have all posed challenges to the traditional accounting framework. At the same time, the state has guided enterprises to actively improve the quality of environmental disclosure by setting up green and low-carbon development funds and giving green credit incentives, forming a policy environment of "hard constraints" and "strong incentives".

2.3 The Evolution and Existing Problems of Environmental Accounting Information Disclosure in Manufacturing Enterprises

Looking back on the past ten years, the environmental disclosure of China's manufacturing enterprises has undergone a process of transformation from scratch, from simple text descriptions to pictures and texts. However, deep-seated problems remain prominent. First, the disclosure caliber is not uniform, due to the lack of mandatory environmental accounting standards, enterprises of different sub-industries and different sizes are more arbitrary in the selection of indicators. Second, the proportion of quantitative information is low, and most enterprises tend to use qualitative words such as "actively responding to national policies" and "significantly reducing emissions", but lack core data support such as energy consumption per unit output value and carbon footprint measurement. Third, the motivation for disclosure is utilitarian, and some enterprises only passively improve the level of disclosure when faced with refinancing or administrative penalty pressure. The existence of these problems reflects that there is still a significant generation gap between the environmental governance capabilities of manufacturing enterprises and the requirements of the "dual carbon" goal.

3. Theoretical Analysis and Research Hypotheses of Influencing Factors

3.1 The Impact Mechanism of Internal Governance Characteristics on the Quality of Disclosure

From the perspective of the enterprise, resource endowment and decision-making mechanism are the basis for determining disclosure behavior. The larger the scale of the enterprise, the higher its attention, and the marginal cost of collecting and processing environmental accounting information is relatively low due to the scale effect, so large enterprises are more likely to provide high-quality information. Profitability is the guarantee of environmental investment, and companies with abundant free cash flow are more capable of green technology transformation and tend to show these results through high-quality reports. The proportion of independent directors reflects the independence of corporate governance, and independent external directors often pay more attention to the long-term compliance risks and social reputation of the enterprise, which can effectively curb the urge of management to conceal negative environmental information[5]. In addition, financial leverage reflects debt pressure, and highly leveraged enterprises may reduce non-mandatory environmental protection expenditures and supporting accounting work due to limited funds due to strict constraints from creditors, resulting in a decline in quality. Based on this, this paper proposes that scale, profitability, and the proportion of independent directors are positively correlated with disclosure quality, while financial leverage is negatively correlated with disclosure quality.

3.2 The Mechanism of the Influence of External Institutional Pressure on the Quality of Disclosure

The external environment constitutes the "constraint boundary" of enterprise decision-making. The intensity of government supervision is the primary factor, and with the normalization of environmental protection inspections, the illegal costs faced by manufacturing enterprises have surged, and this rigid pressure has forced enterprises to standardize environmental accounting processes. The degree of industry competition also plays an important role, and in the highly competitive manufacturing field, environmental performance has become one of the strategies for differentiated competition. In addition, media attention, as a non-institutionalized external supervision, forms a "soft constraint on reputation" on enterprises. When a company is exposed by the media for environmental violations, its market value often shrinks significantly, and in order to restore its reputation, it will repair public trust by improving the transparency of information disclosure[6]. Therefore, this paper proposes that government supervision, industry competition, media attention and disclosure quality are positively correlated.

3.3 Logical Integration and Mediation Effect Analysis of Hypotheses

Internal and external factors are not simply superimposed, but have interactive influences. For example, government regulatory pressure may increase disclosure levels by enhancing environmental management functions within companies. In the context of "dual carbon", the quality of environmental accounting information disclosure is not only the result of a single variable, but also a dynamic balance reached by enterprises under the demands of multiple stakeholders[7]. The hypothesis system constructed in this paper aims to comprehensively cover this complex game process, and construct a three-dimensional causal explanation chain from multiple dimensions such as asset characteristics, profitability level, debt structure, governance mechanism, administrative constraints, and public opinion pressure.

4. Research Design and Empirical Testing Process

4.1 Sample Selection and Data Source Description

This paper selects the A-share manufacturing listed companies in Shanghai and Shenzhen from 2021 to 2025 as the research sample. The reason why 2021 was chosen as the starting point is because it is the first year of "double carbon", and the relevant policy orientation is the clearest. The data processing process is as follows: excluding ST and *ST enterprises with abnormal operations; Samples with missing key financial indicators were eliminated; Winsorize all continuous variables by 1% to eliminate outlier effects. The data sources mainly include: CSMAR database, Wind financial terminal, official websites of environmental protection departments at all levels, and manually collected corporate social responsibility reports.

4.2 Definition of Variables and Construction of Index System

In order to quantify the quality of environmental accounting information disclosure (EDI), this paper uses the content analysis method to empower and score from two levels: monetized information and non-monetized information.

Table 1 Variable categories and definition tables

Variable categories	Variable name	symbol	Definition and measurement method
Interpreted variable	Quality of environmental disclosures	EDI	The total scoring covers 10 dimensions such as environmental protection investment, pollution emission reduction, and green research and development
Explanatory variable 1	Enterprise size	SIZE	The natural logarithm of the total assets of the enterprise
Explanatory variable 2	Profitability	ROE	Return on Equity (Net Profit/Net Assets)
Explanatory variable 3	Financial leverage	LEV	Asset-liability ratio (total liabilities/total assets)
Explanatory variable 4	Proportion of independent directors	INDEP	Number of independent directors/total number of boards
Explanatory variable 5	Government supervision intensity	GOV	Local environmental governance investment amount/regional GDP
Explanatory variable 6	Media attention	MEDIA	The frequency of online media coverage of the company's environmental protection news is naturally logarithmic
Control variables	Enterprise age	AGE	The sample year minus the year of establishment of the enterprise
Control variables	Institutional shareholding ratio	INST	Proportion of total shares held by institutional investors

4.3 Empirical Model Construction and Regression Logic

In this paper, the following multiple linear regression model is established:

$$EDI_{i,t}=\alpha_0+\beta_1 SIZE_{i,t}+\beta_2 ROE_{i,t}+\beta_3 LEV_{i,t}+\beta_4 INDEP_{i,t}+\beta_5 GOV_{i,t}+\beta_6 MEDIA_{i,t}+\gamma Controls_{i,t}+\epsilon_{i,t}$$

In the regression analysis, descriptive statistics are first performed to observe the degree of disclosure quality among manufacturing enterprises. Secondly, Pearson correlation analysis was performed to ensure that there was no serious collinearity problem between the interpreting variables. Finally, the panel fixed effects model is used to regress, and the industry and year effects are controlled, so as to obtain the substantial influence coefficient of each factor on the quality of disclosure.

4.4 Analysis and Discussion of Empirical Results

The regression results showed that the coefficient of enterprise size (SIZE) was significantly positive, which verified that large manufacturing enterprises had a higher quality of environmental disclosure due to their stronger social responsibility awareness and resource mobilization ability. The positive correlation between profitability (ROE) and EDI shows that financial position is the material basis for environmental compliance. It is worth noting that the pass rate of the independent director ratio (INDEP) is extremely high, which suggests that in the "dual carbon" transformation, independent directors with environmental protection professional background or high professionalism have played a key role in correcting the short-sighted behavior of management. Among the external factors, government regulation (GOV) has the highest marginal contribution rate, reflecting that China's manufacturing enterprises are still in the "administrative-driven" disclosure stage.

The salience of MEDIA attention shows that public supervision has become an effective means to make up for the blind spots of administrative supervision.

5. Countermeasures and Guarantees to Improve the Quality of Disclosure under the "Dual Carbon" Goal

5.1 Improve Environmental Accounting Standards and Unified Disclosure Framework

In view of the current problem of fragmentation of disclosure content, the regulatory authorities should formulate accounting standards for the manufacturing industry as soon as possible. The standards should clarify the accounting recognition attributes of carbon emission rights, standardize the classification standards of carbon emission reduction expenditures, and list "energy consumption intensity", "low-carbon R&D investment", and "environmental litigation risk" as mandatory disclosure items. At the same time, a unified digital disclosure platform should be established, requiring enterprises to use structured languages such as XBRL to submit environmental data to improve the comparability and traceability of information.

5.2 Strengthen External Audit Supervision and Market Evaluation Mechanisms

Third-party environmental audit institutions with professional qualifications should be actively cultivated. Currently, relying solely on financial auditors is difficult to identify environmental vulnerabilities in manufacturing processes. By introducing third-party audits, greenwashing can be effectively curbed. At the same time, the China Securities Regulatory Commission and exchanges should optimize the ESG evaluation system and link the quality of environmental information disclosure with the financing authority, listing access, and equity financing costs of enterprises. Establish an "environmental blacklist" system, impose credit restrictions on manufacturing enterprises with seriously low disclosure quality, and use market means to force enterprises to improve themselves.

5.3 Improve the Internal Green Governance Capabilities and Informatization Level of Enterprises

Manufacturing companies should elevate environmental governance to a strategic level and establish an environmental management committee directly led by the board of directors. In internal management, environmental cost accounting (ECA) should be introduced to accurately allocate environmental protection expenditures to each production link, so as to discover the potential points of energy conservation and emission reduction. In addition, enterprises should increase investment in green governance informatization, monitor carbon emission data in real time through IoT sensors, and ensure the authenticity of underlying financial data. Only by realizing "greening" from internal processes can environmental accounting information disclosure obtain living water at the source.

5.4 Cultivate Public Participation Awareness and a Multi-Party Collaborative Regulatory Environment

The government should further disclose environmental enforcement information to provide a basis for media and public scrutiny. By carrying out the selection of green enterprises and publishing the annual report on environmental responsibility in the manufacturing industry, we will create a public opinion atmosphere in which the whole society pays attention to environmental performance. Encourage manufacturing enterprises to cooperate with scientific research institutes to disclose the application prospects of cutting-edge green technologies, so that environmental information disclosure is not only a risk disclosure, but also an opportunity discovery. Through the collaboration of the government, enterprises, market and society, a multi-level environmental accounting information governance ecosystem is built.

6. Conclusion

Based on the background of the "dual carbon" strategy, this paper draws a series of practical conclusions through a systematic investigation of the influencing factors of environmental accounting information disclosure quality of China's manufacturing enterprises. The study finds that the environmental accounting behavior of manufacturing enterprises is the result of the resonance of multiple internal and external factors: strong asset scale and profitability provide the material basis for disclosure, the optimized independent director structure provides the internal governance power, and strong government supervision and active media supervision constitute the external coercive pull. In the second half of the "dual carbon" transformation, the carbon constraints faced by the manufacturing industry will be transformed from "external pressure"

to “internal financial costs”, which requires environmental accounting disclosure to achieve a qualitative change from compliance-oriented to value-creation-oriented.

Although this paper reveals the core causal relationship through empirical research, there is still some room for improvement. For example, due to the highly heterogeneous nature of manufacturing sub-sectors (e.g., chemicals, electronics, machinery), future studies can be conducted horizontally for sub-industries with different pollution attributes. In addition, with the national expansion of the carbon trading market, the impact of carbon price fluctuations on the quality of accounting disclosure will also become a new research hotspot. In short, on the journey to achieve the “dual carbon” goal, true and transparent environmental information is the basis for the harmonious coexistence of green waters and green mountains and golden mountains and silver mountains. Only by continuously optimizing the internal and external factors that affect the quality of disclosure can we truly play the guiding role of accounting in resource allocation and drive China’s manufacturing industry towards a high-quality, sustainable and low-carbon future.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Lei, N., & Wang, X. (2025). From dual-carbon management to green innovation: The synergistic pathways of carbon accounting and carbon footprint in technology enterprises. World Journal of Economics and Business Research, 3(5). https://doi.org/10.61784/WJEBR3071

[2] Chu, X., Yu, M., Chen, Y., et al. (2025). Exploration of issues in carbon accounting services of accounting firms under the “dual carbon” targets. Journal of Social Science and Cultural Development, 1(4). https://doi.org/10.70767/JSSCD.V1I4.519

[3] Babajide, O., Mohammad, A., Adeyemi, K. A., et al. (2024). Determinants and impact of production planning and control accounting techniques on competitiveness of manufacturing companies: A structural equation modelling approach. Accounting Research Journal, 37(4), 453–478. https://doi.org/10.1108/ARJ-01-2023-0014

[4] Li, H. (2024). Risks and preventive measures for manufacturing enterprises in the context of cloud accounting. Applied Mathematics and Nonlinear Sciences, 9(1). https://doi.org/10.2478/AMNS.2023.1.00399

[5] Chowdary, V. B., & Harrison, S. (2024). Lean practices: Assessing the relationship among manufacturing and service firms. International Journal of Industrial and Systems Engineering, 47(4), 421–451. https://doi.org/10.1504/IJISE.2024.140340

[6] Qais, S. Y., Munther, N. A. B., & S, S. A. (2023). The quality of cost accounting systems in manufacturing firms: A literature review. Cogent Business & Management, 10(1). https://doi.org/10.1080/23311975.2023.2209980

[7] Hien, T. N. (2022). Factors affecting the implementation of environmental management accounting: A case study of pulp and paper manufacturing enterprises in Vietnam. Cogent Business & Management, 9(1). https://doi.org/10.1080/23311975.2022.2141089

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Optimization of Professional Master Enrollment Promotion Model Based on Lasswell's "5W" Model

Yang Wang*, Pengzhi Kong

MBA School, Shandong University of Finance and Economics, Jinan, Shandong, 250014, China

Corresponding author: Wang Yang, 20188483@sdufe.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: China's postgraduate education is moving towards high-quality development with increasingly fierce enrollment competition. Despite the consecutive decline in applicant numbers, the enrollment quota and proportion of professional degree postgraduates are on the rise, creating critical challenges for their enrollment and emphasizing the importance of effective promotion. As a targeted communication activity, enrollment promotion inherently aligns with Lasswell's "5W" model. This study applies communication theory and qualitative analysis to explore the five core elements of professional degree postgraduate enrollment promotion, and constructs an optimized 5W-based promotion framework.

It proposes building a four-dimensional communication team of admissions offices, supervisors, current students and alumni to boost credibility; designing candidate-centric promotional content highlighting practical and career advantages to create a "phenomenal" communication effect; integrating online-offline media channels to form a multi-dimensional communication matrix; utilizing big data to create precise candidate portraits for targeted enrollment; and establishing a comprehensive feedback mechanism to dynamically adjust strategies. This research provides practical references for universities to enhance their professional degree postgraduate enrollment promotion amid the complex and volatile current landscape.

Keywords: Lasswell's "5W" Model; Professional Degree; Postgraduate Enrollment Expansion; Decline in Application Numbers; Enrollment Promotion Model

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1413

1. Introduction

According to the "Statistical Communique of the People's Republic of China on National Economic and Social Development in 2022" released by the National Bureau of Statistics, at the end of 2022, the national population experienced negative growth for the first time since 1962, with a natural population growth rate of -0.60‰. [1] At the same time, significant changes have also taken place in the postgraduate application situation in the field of education. In 2025, the number of national master's degree applicants was 3.88 million, a decrease of 500,000 compared with the previous year, a decline of about 11.4%. [2] This change profoundly reflects that China is undergoing a profound transformation in terms of population structure and educational needs. Against the background of negative population growth, the supply and demand relationship in the social labor market has changed, putting forward new requirements for the quality and structure of talents. All walks of life have an increasingly strong demand for applied talents with high professional capabilities and outstanding practical abilities. In

this trend, professional master's degree education, which focuses on cultivating students' professional skills and practical application capabilities, has become an important way to meet the social demand for applied talents. In the 2025 master's degree unified examination enrollment plan, the planned enrollment of professional master's degree students was 602,000, accounting for 68.99% of the total enrollment plan, far exceeding the number of academic master's degree enrollments. The expansion of professional degree postgraduate enrollment has become the mainstream trend.

The expansion of enrollment scale has intensified the competition among universities for professional degree postgraduate students. However, the current enrollment promotion work is facing many challenges. To address the challenges and improve the scientificity and effectiveness of enrollment promotion, it is particularly necessary to draw theories and experiences from relevant disciplinary fields. Lasswell's "5W" model in communication science provides a comprehensive and systematic analytical framework for enrollment promotion work. By in-depth analysis of the key elements in enrollment promotion activities, such as communicators, communication content, communication channels, audiences, and communication effects, it can help universities optimize promotion strategies, accurately locate target audiences, improve the quality and attractiveness of promotion content, select appropriate communication channels, thereby enhancing the effect and influence of enrollment promotion, maximizing the value of enrollment promotion work, and promoting the high-quality development of professional degree postgraduate education.

2. Theoretical Basis

Lasswell, one of the four founding fathers of communication studies, proposed the "5W" model in his work The Structure and Function of Communication in Society. This model comprises five core elements: Who (the communicator), Says What (the message), In Which Channel (the medium), To Whom (the audience), and With What Effect (the effect). [3] Widely applied in the field of mass communication, this model provides a fundamental framework for communication research. Despite its certain limitations—such as neglecting the feedback loop and the impact of the communication environment—it undoubtedly holds significant value in analyzing specific communication activities.

Li Bin argued that the 5W model introduces five major elements that have shaped the development of communication, covering nearly all content within the field and forming the theoretical framework of the entire discipline. [4] Zhang Guoliang emphasized two key values of the model: first, it represents the first detailed dissection of the communication process in communication studies; second, it was the first to propose five major research areas in communication, offering new perspectives and pathways for subsequent in-depth and systematic exploration of communication theories. The advent of the 5W theory established a comprehensive framework, laying a solid foundation for the study of communication theories. [5]

Scholars have also critically examined the limitations of the 5W model. Lasswell's remarkable contribution to communication studies lies in his provision of a clear direction for research, enabling scholars to identify focal points of investigation. For nearly half a century, the field of communication has focused on exploring the fundamental elements of the communication process, and the emergence of the 5W model has laid a profound groundwork for its long-term development. Lasswell's communication theory has served as a beacon for the discipline over more than half a century.[6]

Enrollment promotion is essentially a form of communicative behavior. Its core objective is to attract potential students, foster their interest in the institution, and ultimately influence their decision to apply. This process involves shaping students' perceptions and guiding their decision-making. Currently, there exists a substantial body of research on enrollment promotion from a communication studies perspective in China.

For instance, drawing on Lasswell's "5W" model and combined with Augmented Reality (AR) technology, Li Han (2022) took Harbin Institute of Technology as a case study to design an AR enrollment promotion brochure. Through demand analysis, design and development, and evaluation, this approach addressed issues such as one-way information dissemination and homogenization inherent in traditional materials, effectively communicating enrollment information and the university's spirit. This research provided a reference for innovation in university enrollment promotion, facilitating the attraction of high-quality applicants. [7]

Wang Jin (2021) applied the "5W" model to analyze enrollment promotion at the University of Science and Technology Beijing, analyzing the current situation through each communicative element and proposing improvement strategies. This

study offered insights for higher education institutions to enhance enrollment effectiveness and improve the quality of undergraduate applicants. [8]

Lu Wei et al. (2020) utilized the "5W" model to examine enrollment promotion at Southwest Jiaotong University, exploring strategic shifts and existing issues in the new media era, and putting forward optimization suggestions. Their work provided a reference for innovation in university enrollment promotion. [9]

Wu Yanzhe and Liang Rong (2015) analyzed enrollment promotion in vocational colleges using the "5W" model, dissected the process elements, and constructed corresponding strategies. This research provided theoretical support and practical reference for vocational colleges to enhance the effectiveness of their enrollment promotion efforts. [10]

Zhang Lina and Li Yan (2019) employed the "5W" model to analyze enrollment promotion for international students in Chinese higher education institutions. After clarifying the current situation and problems, they designed effective strategies to improve the accuracy and effectiveness of student recruitment. [11]

Therefore, based on the characteristics of enrollment promotion activities and existing preliminary research, this study introduces Lasswell's model. It aims to conduct an in-depth analysis of the various elements within professional master's enrollment promotion using this theoretical framework, optimize promotion strategies, enhance the effectiveness and targeting of enrollment efforts, better achieve enrollment objectives, and promote the high-quality development of professional master's education.

3. Analysis of Problems in Professional Master's Enrollment Promotion Based on the "5W" Model

The "5W" communication model proposed by Lasswell provides a systematic analytical framework for decomposing the enrollment promotion work of professional master's degree postgraduates. Essentially, enrollment promotion is a communication activity that accurately reaches target groups and conveys core values, and its effectiveness depends on the coordinated efforts of five links: communicators, messages, media, audiences, and effect feedback. Many current pain points in professional master's enrollment promotion correspond exactly to the shortcomings of each dimension of the "5W" model, and the problems in each link are interrelated and influence each other, forming a closed-loop bottleneck that restricts the promotion efficiency. The following specifically analyzes the prominent problems in enrollment promotion work in combination with each core dimension of the "5W" model.

3.1 Single Communicator Structure: Failure to Form a United Propaganda Force

In the traditional postgraduate enrollment process, the communicators exhibit a homogenized structure, failing to effectively integrate resources and consolidate a united propaganda force. In practice, the primary task of information dissemination in postgraduate enrollment mostly falls on the university's admissions office. This department shoulders key responsibilities such as formulating enrollment policies, releasing enrollment information, and planning promotional activities, and its work possesses a certain degree of professionalism and targeting. However, this communication model, which relies entirely on a single department, greatly restricts the breadth and depth of enrollment information dissemination. While the admissions office is professional, it lacks effective collaboration with other departments (colleges, research teams, alumni associations, etc.), leading to a flattened information transmission process. This prevents a comprehensive and multi-level display of the university's enrollment advantages and characteristics. Especially since professional master's degree education focuses on practical application, the involvement of industry experts and corporate partners should be emphasized. A single enrollment subject makes it difficult to gather diverse forces for enrollment promotion, resulting in a weak propaganda offensive and significantly reduced efficiency.

3.2 Outdated and Vague Messages: Severe Homogenization of Enrollment Promotion Content

Although enrollment promotion materials have evolved from print media to printed matter and then to electronic publications, their content remains highly homogenized, lacking uniqueness and differentiation.[12] Currently, university profiles, discipline classifications, faculty rosters, academic achievements, and enrollment policies still form the core promotional content of most universities, creating a stereotyped propaganda routine. The presentation is often perfunctory and uniform, failing to highlight key points and overlooking the unique concepts and distinct differences among universities in professional master's

degree cultivation. Consequently, candidates struggle to accurately distinguish the unique charm and competitive advantages of each university and its programs amidst the massive volume of enrollment information. Furthermore, the lack of innovation in promotional content—being outdated and vague—leads to extensive redundant information, failing to leave a lasting impression on candidates and yielding unsatisfactory promotional results.[13]

3.3 Insufficient Media Innovation: Limited Effectiveness of Enrollment Promotion Methods

At present, many domestic professional master's degree enrollment institutions still rely on traditional methods as their primary means of promotion, such as releasing regular enrollment updates, setting up enrollment inquiry hotlines, distributing printed enrollment brochures, and conducting off-site enrollment lectures in student source areas. This enrollment promotion pathway, centered on traditional models, has numerous drawbacks that hinder the efficient progress of enrollment work. On the one hand, the efficiency of information transmission is low. Printed enrollment brochures and materials incur high production and distribution costs with limited dissemination effects and slow transmission speeds. Enrollment updates posted on official websites mainly reach candidates with active application intentions, making it difficult to effectively reach a broader target candidate group. On the other hand, there is a lack of interactivity. Traditional enrollment promotion models are mostly one-way communication, lacking interaction with potential candidates. This makes it challenging for candidates to fully understand the university and for the university to promptly address their questions and needs. The over-reliance on traditional, rigid promotion methods with insufficient innovation fails to meet the diverse application needs of candidates from different groups.

3.4 Undifferentiated Audience Targeting: Lack of Precision in Enrollment Promotion Objects

In the field of professional master's degree enrollment promotion, the accuracy of audience positioning plays a crucial role in achieving promotional effects. However, the current situation is worrying: enrollment promotion exhibits obvious undifferentiation in terms of the audience, lacking meticulous precise positioning and in-depth analysis of potential student groups. When conducting enrollment promotion activities, institutions often adopt a "wide-net" strategy. While they can cast a wide net, they fail to focus on key targets, resulting in high costs and low efficiency.[12] In practice, institutions usually release a large number of generalized promotional materials through various enrollment promotion channels. These materials lack targeting, failing to fully consider the differentiated needs of different potential student groups. Due to the absence of precise positioning, it is difficult for universities to highlight their own advantages amidst numerous enrollment promotions, or to showcase the unique value of their professional master's degree education to potential students. For candidates, facing a massive volume of homogenized promotional information, it is challenging to screen out the universities and programs that truly meet their own requirements and expectations. This leads to low contact efficiency with the target group, an imbalance in the input-output ratio of enrollment promotion work, and a situation of high cost and low efficiency.[14]

3.5 One-way Effect and Feedback Process: Absence of a Feedback and Evaluation Mechanism for Enrollment Promotion

Effective enrollment promotion requires smooth feedback channels to promptly understand the needs and feedback of potential candidates. Nevertheless, current enrollment promotion practices in domestic universities lack systematic and comprehensive summary and evaluation of student source quality. Horizontally, it is difficult to reflect the gaps compared with other postgraduate enrollment institutions at the same level; vertically, there is also a lack of comparative evaluation of the quality of student sources in the institution over the years, and a dynamic adjustment mechanism for enrollment promotion forms and methods has not been formed.[15] During the enrollment promotion process, universities only focus on one-way information output, and there is a serious lack of interaction and communication mechanisms with potential candidates, making it difficult to timely grasp the actual needs of candidates and their feedback on promotional content. Evaluating the effectiveness of enrollment promotion is a key link in improving promotional strategies. The lack of clear and definite evaluation criteria leads to great subjectivity and uncertainty in evaluation results, making it difficult to adjust and optimize promotional strategies. On the one hand, it fails to provide a solid and reliable basis for the scientific adjustment of promotional strategies, causing enrollment promotion work to linger at a low level and making it difficult to achieve a qualitative breakthrough. On the other hand, it is impossible to conduct fair and effective assessment and incentives for the performance of different departments or

individuals in enrollment promotion work, making it difficult to fully mobilize the enthusiasm and creativity of all parties. Ultimately, this leads to low efficiency of enrollment promotion work, failure to give full play to its due effectiveness, and also affects the selection of high-quality student sources to a certain extent.

4. Construction of Enrollment Mode Based on Lasswell's "5W" Model

In the practice of enrollment promotion for professional degree postgraduate students, Lasswell's "5W" model provides a solid theoretical foundation and clearly outlines the key context of the communication process. Starting with an in-depth analysis of the five basic elements of the communication process—communicator, message, medium, audience and effect—this paper accurately identifies the current situation and problems of each element, and then formulates targeted strategies to optimize and upgrade the enrollment promotion mode for professional degree postgraduate students.

4.1 Optimizing the Communicator Element to Form a Four-Dimensional Promotion Team of "Admissions Office, Supervisors, Current Students and Alumni"

The training of professional degree postgraduate students focuses on practical application and professional literacy, so their enrollment promotion should meet the core needs of candidates and optimize the construction of the communicator team. Combined with the practices of universities in many provinces and cities, the establishment of a four-dimensional promotion team of "Admissions Office–Supervisors–Current Students–Alumni" can transform promotion from "one-way indoctrination" to "two-way interaction" and improve its effectiveness.

As the overall coordinator, the Admissions Office is responsible for policy release, platform operation and team support. At Zhejiang University, the university and school-level admissions offices collaborate to form promotion teams by region, compile promotion manuals and conduct regular volunteer training, solving the problem of fragmented promotion efforts.

The supervisor promotion team fully satisfies the credibility effect in communication studies: the higher the credibility, the greater the persuasive effect, and vice versa [10]82. South China University of Technology has established a dual-supervisor team of "industry + academia", inviting industrial supervisors from enterprises such as Huawei to give lectures based on their experience, while academic supervisors interpret training programs, providing all-round guidance for promotion activities.

The current student team conveys real campus information and narrows the distance with candidates. Shandong University and other universities have set up enrollment ambassador teams, grouped by region, to share exam preparation and practical experience through live broadcasts and consultation groups, alleviating candidates' information anxiety.

The alumni team broadens promotion channels by virtue of resource advantages. Shandong University of Finance and Economics relies on alumni associations to hold special lectures at alumni enterprises in many regions, expands promotion through enterprise cooperation, and builds a bridge for employment connection to achieve a virtuous cycle.

In summary, through division of labor and collaboration, the four-dimensional team meets candidates' information needs regarding future learning environment, research directions, employment status and other aspects, helping candidates form a true, comprehensive and objective perception [16]. Cases from the four universities fully prove its feasibility, and its operation mode can provide a reference for other universities to effectively solve the problem of vague and general promotion.

4.2 Highlighting Message Advantages, Candidate-Oriented, and Creating a "Phenomenal" Effect

From the perspective of messages, the content of professional degree postgraduate enrollment should be accurately positioned around candidates' needs and the unique orientation of professional degree education. On the one hand, in content construction, the practical-oriented characteristics and career development advantages should be highlighted. Practical courses in the professional degree curriculum system should be introduced in detail, including case teaching closely integrated with industry practice, project-driven learning, enterprise internships and training, enabling candidates to clearly understand how the major transforms theoretical knowledge into practical operational capabilities through practical teaching, thereby enhancing their future workplace competitiveness. Attention should be paid to demonstrating the career development advantages of professional degrees, providing detailed employment data and destination information such as the employment ratio of graduates in different industries and enterprises, average salary levels and career promotion paths, outlining a clear career development blueprint for candidates.

On the other hand, from a design perspective, adhering to a student-oriented approach, we should accurately position can-

didates' needs, gain an in-depth understanding of their exam preparation links and key concerns such as question difficulty, exam skills and employment prospects, so as to design targeted and practical promotional materials. For example, producing exam preparation guides, mock tests and experience-sharing resources helps candidates better prepare for exams and plan their future, subtly making them recognize the university brand and optimizing promotion effects. In-depth exploration of outstanding deeds of current students, such as scientific research achievements, social practice and voluntary services, should be widely publicized through press releases, videos and other forms. Current students and alumni take the initiative to release and spread relevant information, becoming "information hubs" [15]20, radiating to candidates' social circles through mainstream online media platforms such as WeChat Moments, forming a point-to-area "phenomenal" effect and greatly improving the influence and popularity of the major.

4.3 Constructing Diversified Enrollment Media Channels, Building a Promotion Media Matrix, and Integrating Online and Offline Promotion

First, develop distinctive enrollment channels for precise connection with target students. Industry exhibitions are important windows for displaying practical achievements. Universities can intuitively present students' practical works at exhibitions to demonstrate the effectiveness of professional practical training. For instance, the Hongqi College of Northeast Normal University displays practical project achievements of professional degree postgraduate students and school-enterprise cooperative training works at industry exhibitions related to smart vehicles, attracting the attention of participating enterprises and potential candidates, and simultaneously conducting on-site consultations with cooperative enterprises such as FAW, realizing two-way empowerment of enrollment promotion and industry-education integration. Meanwhile, deepen school-enterprise cooperation and expand internal recommendation channels relying on alumni enterprises. For example, China Jiliang University cooperates with alumni enterprises including Zhejiang China Commodity City Group Co., Ltd., using internal enterprise communication networks to accurately push enrollment information to in-service personnel with career promotion needs, greatly improving promotion pertinence. Southwest University carries out order-based training for professional master students with Changan Automobile, selecting candidates meeting job requirements through internal enterprise recommendation channels to form a closed loop of "enrollment–training–employment".

Second, integrate resources to build a powerful promotion matrix and expand communication coverage. With official new media as the core, a promotion media matrix should be established to improve promotion quality and efficiency [17]. For example, China Jiliang University takes its postgraduate enrollment WeChat official account as the carrier, links official accounts of secondary colleges to push enrollment updates and major characteristics in real time, and simultaneously cooperates with mainstream platforms such as the National Postgraduate Enrollment Network and China Education Online for retransmission and communication. It integrates and promotes QR codes of official accounts, video accounts, official websites and consultation groups, integrating them into university brand building to achieve multi-channel coordinated efforts and effectively improve promotion accuracy.

Finally, promote in-depth integration of online and offline promotion to strengthen practical implementation. Offline, trust is built through interaction, and face-to-face communication with candidates is realized through lectures, course trials and other activities. For example, Shandong University of Finance and Economics conducts offline lectures in cities and enterprises with high-quality student sources to enhance candidates' intuitive experience. Online, coverage is broadened through live broadcasts, cloud consultations and other forms to bridge information gaps. Through the linkage of online live lectures and offline promotion meetings, Shandong University of Finance and Economics replays highlights of offline activities online and guides online consultations to offline participation, forming a linkage mode of "offline as the foundation, online as the wing", injecting sustained impetus into enrollment work.

4.4 Accurately Positioning the Audience, Formulating Candidate Portraits, and Implementing Targeted Enrollment

To accurately identify the needs of professional degree postgraduate candidates, from the audience perspective, enrollment universities need to conduct in-depth group and individual analyses. Group analysis is used to grasp overall trends, such as the age, professional background, educational level and industry distribution of target candidates, helping us understand

their main demands such as career development bottlenecks, management ability improvement or academic background optimization. Individual analysis focuses on exploring each candidate's personalized needs, including application doubts such as university selection, curriculum setting and employment prospects, as well as exam preparation difficulties such as time management, knowledge mastery and interview skills.

By collecting and analyzing data of registered candidates, we can customize audience portraits based on analytical results. Portrait technology, from the specific perspective of marketing needs, outlines users' behavioral characteristics and consumption intentions to present clearer and more intuitive user needs [18]. For example, a typical portrait of an in-service candidate is a 30-year-old male programmer with 5 years of work experience, pursuing a professional degree to improve technical management capabilities. Such candidates have core demands for mastering knowledge and skills and improving management levels, as well as pain points such as work-study conflicts. This strategy precisely connects university advantages with candidates' demands, enhancing the attractiveness and recognition of enrollment information among target groups through differentiated and customized promotional content.

Through accurate portrait analysis, enrollment promotion can tap into candidates' potential needs, realize efficient resource allocation in the selection of promotion channels, writing of promotional copy and planning of enrollment activities, improve the input-output ratio of enrollment promotion, promote the refinement and scientization of enrollment work, and help universities expand the scale and improve the quality of student sources in the fiercely competitive professional master enrollment market.

4.5 Establishing an Enrollment Effect Feedback Mechanism to Promote Innovation in Enrollment Promotion

From the effect perspective, professional degree postgraduate enrollment promotion should establish a horizontal and vertical comparative feedback mechanism to timely analyze, compare and summarize annual enrollment promotion work. Horizontally, compare the enrollment promotion practices of similar universities, face up to the university's brand competitiveness, analyze the quality of student sources, and identify its competitive position in student recruitment. Vertically, compare the effectiveness of promotion channels and candidate conversion rates over the years. With the help of background data of integrated media platforms, statistics and summaries of reading volume of WeChat official account articles and click volume of video accounts for enrollment promotion should be conducted, continuously tracking and analyzing the promotion effects of different channels and adjusting delivery strategies timely based on data feedback. For example, if the interaction rate of promotional content on a social media platform is low, it may be necessary to re-examine the content form or delivery time, or consider switching to other channels more suitable for the target audience. Meanwhile, flexibly adjust channel combinations according to different stages of the enrollment cycle. In the early stage of enrollment, a wide range of channels can be used to increase visibility; in the sprint stage of registration, focus on several key channels where target audiences are most active for precise push and reminders.

In addition, a closed loop should be formed from top-level design, plan formulation, teacher and student selection, training, activity implementation to assessment and evaluation, establishing a sound assessment mechanism to enhance the stability of the enrollment promotion team. [19] Qingdao Agricultural University has accumulated mature experience in this regard, deeply integrating the enrollment promotion feedback mechanism with the assessment mechanism to form a virtuous cycle of "data feedback–strategy adjustment–assessment and evaluation–optimization and improvement". The university regularly summarizes and analyzes effect data of various promotion channels to form feedback reports, timely adjusts inefficient promotion strategies, and directly links promotion performance with the performance assessment of the promotion team, commending and rewarding outstanding teachers and students, and conducting interviews and guidance for underperforming personnel, continuously improving the work quality of the promotion team.

5. Conclusion

This paper analyzes the application of Lasswell's "5W" model in enrollment promotion strategies for professional degree postgraduate students, and creatively proposes a series of strategies and methods aimed at improving the efficiency and quality of enrollment promotion. By constructing a four-dimensional promotion framework covering "Admissions Office, supervi-

sors, current students and alumni", the composition of the enrollment team is optimized, ensuring the comprehensiveness and depth of promotion forces. In terms of content strategy, focusing on highlighting practical and professional advantages and closely centering on candidates' concerns, it stimulates a "phenomenal" communication effect, achieving precise reach and wide resonance of promotional content. In terms of communication channels, a three-dimensional and diversified enrollment channel system is constructed, and a multi-dimensional promotion media matrix is built to broaden the breadth and depth of promotion. For the audience, precise marketing and targeted enrollment are realized through accurately portraying portraits of target candidates. In addition, the establishment and implementation of horizontal and vertical enrollment promotion feedback mechanisms continuously drive the innovation and upgrading of enrollment promotion, attract outstanding student sources, and further promote the high-quality and connotative development of professional degree postgraduate education in universities.

In summary, to improve the enrollment quality of professional degree postgraduate students, universities should make full use of their own teaching and research advantages, gain an in-depth insight into candidates' needs and market dynamics, actively integrate internal and external resources, and form an efficient, coordinated and precisely implemented enrollment promotion system.

Funding

This paper is part of the achievements of the 2025 Postgraduate Education and Teaching Reform Project of Shandong University of Finance and Economics, "Research on the Optimization of MBA Enrollment Promotion Path Based on Lasswell's '5W' Model — A Case Study of Shandong University of Finance and Economics" (No. jxgg2025007).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] National Bureau of Statistics. (2023, February 28). *Statistical communiqué of the People's Republic of China on the 2022 national economic and social development*. Retrieved November 26, 2024, from https://www.stats.gov.cn/sj/zxfb/202302/t20230228_1919011.html

[2] Ministry of Education. (2024, November 21). *Ministry of Education deploys security work for 2025 national postgraduate admission examination*. Retrieved November 26, 2024, from http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/moe_1485/202411/t20241121_1163982.html

[3] Duan, P. (2020). *Fundamentals of communication: History, framework and extension*. Communication University of China Press.

[4] Li, B. (1993). *Introduction to communication studies*. Xinhua Publishing House.

[5] Zhang, G. L. (2009). *Principles of communication*. Fudan University Press.

[6] Lin, Z. D. (1996). A new theoretical research framework of communication studies. *Journalism Quarterly*, 32.

[7] Li, H. (2022). *Research on the design of university AR enrollment brochure based on 5W communication model* [Master's thesis]. Harbin Institute of Technology.

[8] Wang, J. (2021). University enrollment promotion strategies in the new era based on the 5W model: A case study of University of Science and Technology Beijing. *Education Media Research*, (02), 40–42.

[9] Lu, W., Wang, A. Q., Fang, Y. R., et al. (2020). Analysis of university enrollment promotion strategies in the new media era based on 5W model: A case study of Southwest Jiaotong University. *Public Communication of Science & Technology*, 12(14), 26–28.

[10] Wu, Y. Z., & Liang, R. (2015). Exploration on the effectiveness of enrollment promotion in higher vocational colleges based on Lasswell's 5W model. *China Adult Education*, (10), 80–82.

[11] Zhang, L. N., & Li, Y. (2019). Research on enrollment promotion strategies for international degree students in universities based on Lasswell's 5W model: A case study of a western university. *The Science Education Article Collects*, (16), 17–18.

[12] Ni, R., Du, J., & Li, W. T. (2020). Exploration on precision enrollment promotion for postgraduates in the information age: A case study of University of Science and Technology of China. *Academic Degrees & Graduate Education*, (06), 44–48.

[13] Su, N., Shi, J. J., Lv, K., et al. (2023). Thoughts and exploration on the innovation of postgraduate enrollment promotion mode in the new era. *Western Journal*, (03), 103–106.

[14] Du, M. M. (2024). Research on the optimization of university enrollment promotion strategies from the perspective of digital transformation. *Journal of Jilin Radio and TV University*, (03), 128–130.

[15] Xie, Y. H., Tang, H., & Tang, H. (2023). Research on enrollment promotion strategies for postgraduates in universities in the new era. *University*, (28), 17–20.

[16] Cai, J. S. (2022). Research on postgraduate enrollment work and promotion strategies from a multidisciplinary perspective. *Teaching of Forestry Region*, (12), 35–38.

[17] Shi, G. (2023). Research on the application of new media matrix in university enrollment promotion. *Journal of News Research*, 14(09), 144–146.

[18] Zhang, W. Q., & Zhao, S. (2021). On the application of Douyin short videos in enrollment promotion of higher vocational colleges. *Vocational Education in Machinery*, (11), 1–7.

[19] Yang, T., & Liu, R. (2021). Exploration and thinking on university enrollment promotion management mechanism under the grid system. *Journal of Yancheng Institute of Technology (Social Science Edition)*, 34(06), 96–99.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

The Digital Economy "Empowers" Carbon Productivity—Empirical Evidence from Chinese Cities

Yu Zhang*

Xi'an Shiyou University School of Economics and Management, Xi'an, Shaanxi, 710065, China

Corresponding author: Yu Zhang, y335856@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: In the context of building a new development pattern, the digital economy plays an increasingly prominent role in promoting the "dual carbon" goal. Based on the urban panel data from 2011 to 2022, this paper deeply explores the impact and mechanism of digital economy on urban carbon productivity. The results show that: (1) The digital economy can promote the improvement of urban carbon productivity, which is still valid after a series of robustness tests. (2) From the perspective of mechanism path, the digital economy can improve urban carbon productivity through two channels: promoting technological innovation and industrial structure upgrading. (3) From the heterogeneity analysis, it can be seen that the promotion effect of digital economy on urban carbon productivity is more obvious in non-resource cities and peripheral cities. This study provides empirical evidence and policy reference for the digital economy to promote green development and achieve the "dual carbon" goal.

Keywords: Digital Economy; Carbon Productivity; Digital Infrastructure

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1419

1. Introduction

Global climate change has become a common challenge facing human society. Reducing carbon emissions and improving carbon productivity have become the core issues to achieve sustainable development. As the largest developing country, China has promised to achieve peak carbon dioxide emissions by 2030 and carbon neutrality by 2060, which puts forward higher requirements for the green transformation of economy and society. The digital economy is regarded as an important engine to promote the improvement of carbon productivity by virtue of its ability of technology penetration and resource optimal allocation. For example, the National Big Data Comprehensive Test Area adopted green technological innovation[1]. Robot application and industrial structure optimization[2], significantly reducing the growth rate of urban carbon emissions and increasing carbon productivity. The digitalization of rural areas is achieved through technological innovation[3], scale of operation and non-agricultural employment and other channels[4], has a positive impact on agricultural total factor carbon productivity. However, the existing research focuses more on economics and management perspectives, and less on the issues of cost accounting, environmental accounting information disclosure and resource allocation efficiency optimization in the digital transformation from the accounting perspective. It is difficult to fully reveal the micro-mechanism of the correlation between digitalization and carbon productivity.

Existing research shows that digitalization mainly enhances carbon productivity through the effects of technological

innovation, industrial structure optimization and resource allocation. For example, the development of the Internet significantly increases carbon productivity in the mid-term of industrialization by improving the efficiency of resource allocation and in the late stage by promoting technological progress[3]; Digital industry collaborative innovation promotes total factor carbon productivity by optimizing the allocation of capital and human factors[5]. In addition, the platform economy plays a "cloak effect" in the digital construction, i.e. it may increase carbon emissions due to infrastructure investment in the short term, but it will suppress emissions through technology diffusion and industrial upgrading in the long term[6]. These studies provide theoretical support for understanding the digital carbon emission reduction path, but lack attention to accounting issues such as cost accounting and carbon asset measurement.

2. Literature Review

Under the background of the global "double carbon" target, digital transformation has become the core engine to improve carbon productivity. The National Big Data Comprehensive Test Area policy significantly reduces the growth rate of urban carbon emissions and increases the growth rate of carbon productivity through three paths: green technology innovation, extensive application of robots and optimization of industrial structure. For example, the pilot area promoted the green transformation of high-energy consuming industries by integrating data resources and promoting industrial internet applications, resulting in a 7.09% decrease in the growth rate of carbon emissions and a 13.84% increase in carbon productivity[7]. Rural digitalization improves agricultural total factor carbon productivity through technology diffusion effect. For example, panel data of 30 provinces in China show that for every 1% increase in rural digitalization level, agricultural total factor carbon productivity will increase by 0.629%, and this effect will be achieved through technology innovation, scale operation and non-agricultural employment[8]. In addition, collaborative innovation in the digital industry promotes total factor carbon productivity more significantly in the eastern region and regions with higher levels of new infrastructure by optimizing the allocation of capital and human factors, such as improving capital mobility and attracting high-quality talents[9].

In terms of technological innovation and efficiency improvement, digitalization directly improves carbon productivity by promoting green technological innovation. For example, the development of the Internet has a positive impact on carbon productivity by improving the efficiency of resource allocation in the middle stage of industrialization and promoting technological progress in the later stage, with the effect coefficient of technological progress reaching 0.0206 in the latter stage of industrialization[10]. The number of green patent applications and green total factor productivity in the National Big Data Pilot Zone have significantly increased, which proves the mediating role of technological innovation[11]. In terms of industrial structure optimization and agglomeration effect, digitalization enhances carbon productivity through rationalization, agglomeration and digitalization transformation of industrial structure. For example, the collaborative innovation of digital industry promotes the collaborative agglomeration of manufacturing industry and producer service industry. The decrease of theil index indicates that the industrial structure is more reasonable[12]. The digitalization of rural areas can reduce the energy consumption and transaction cost of logistics and realize the coordination of virtual agglomeration and geographical agglomeration by promoting industrial clusters such as "Taobao Village"[13]. In terms of spatial spillover and platform economic effect, the impact of digitalization has spatial spillover. The dynamic spatial dobbin model shows that the rural digitalization not only improves the local agricultural total factor carbon productivity, but also has a positive indirect effect on the neighboring regions through the technology diffusion. the short-term and long-term indirect effects are significantly 0.946 and 1.303 respectively[14]. The platform economy has shown a "masking effect" in the digital construction: in the short term, the platform infrastructure investment may increase carbon emissions, but in the long term, emissions are suppressed through the circulation of data elements and industrial upgrading. after the implementation of the "two-carbon" strategy, the masking effect of the platform economy is weakened, and the positive impact of digitization is enhanced[15].

The effect of digitalization has significant regional heterogeneity. Due to the perfect digital infrastructure and intensive innovation elements in the eastern region, the promotion effect of digital industry collaborative innovation on carbon productivity is significantly higher than that in the central and western regions[16]. From the perspective of the evolution of industrialization stage, the impact of internet development on carbon productivity is "double threshold effect". In the early

stage of industrialization, there was a negative impact due to inadequate technological adaptation, while in the middle and later stages, carbon productivity was significantly improved through resource allocation and technological progress[17]. The new infrastructure level and financial development level adjust the digital effect. The synergetic innovation of digital industry in regions with higher level of new infrastructure will promote carbon productivity by three times as much as in regions with lower level[18]; There is a "threshold effect" in the development level of rural finance. Only in areas with high financial level, the rural digitalization has a significant effect on the improvement of carbon productivity[19]. In addition, after the implementation of the "double carbon" strategy, the masking effect of platform economy is weakened and the direct effect of digitalization on carbon productivity is enhanced[20].

3. Theoretical Analysis and Research Hypotheses

The digital economy drives the improvement of carbon productivity through four dimensions of technology, structure, efficiency and space. Its mechanism can be summarized as the progressive path of "technological innovation breakthrough-reshaping industrial structure-optimizing resource allocation-spatial collaborative diffusion".

The Impact of Digital Economy on Carbon Productivity From a Theoretical Perspective, Digital Economy integrates data resources through the effect of technological innovation to drive research and development and application of green technologies[1]. For example, the National Big Data Comprehensive Test Area promoted the green transformation of energy-intensive industries through the industrial internet platform, increasing the green patent application volume by 0.9197% and the green total factor productivity by 6.82%; Relying on the effect of industrial structure optimization, digital technology permeates into the traditional manufacturing industry to promote industrial synergy. As non-resource-based cities have higher elasticity of industrial structure, the effect of digital economy on carbon productivity is significantly higher than that of resource-based cities[11]; Through the effect of resource allocation efficiency, Digital inclusive finance alleviates the financing constraint of green enterprises. Each 1% increase in capital factor allocation efficiency can drive the growth of carbon productivity by 0.679%. At the same time, rural digitalization transfers labor to low-carbon industries through non-agricultural employment channels, and non-agricultural income is increased by 0.235%[13]; With the help of spatial spillover effect, the short-term indirect effect of rural digitalization on carbon productivity in neighboring regions reaches 0.946. In the eastern region, due to the improvement of new infrastructure and the concentration of innovation elements, the promotion effect of digital industry collaborative innovation on carbon productivity is more significant. Based on this, hypothesis 1 is proposed.

Hypothesis 1: The digital economy significantly increases carbon productivity.

The digital economy accelerates the research and application of green technologies through the integration of data elements and the innovation of ecological construction. Big Data, cloud computing and other technologies have reduced information asymmetry and pushed interdisciplinary green innovation projects to the ground. For example, the National Big Data Comprehensive Test Area has optimized production processes through industrial internet platforms[5], increased the green patent application volume by 0.9197% and green total factor productivity by 6.82%[15]. At the same time, the extensive application of robot technology[1], such as industrial robot installation density increase, directly replace the high energy consumption production link, reduce carbon emissions per unit output[5]. Based on this, hypothesis 2 is proposed.

Hypothesis 2: Technological Innovation Is the Mediating Path of Digital Economy Affecting Carbon Productivity

Digital Economy Promotes Green Transformation of Traditional Industries and Agglomeration of Emerging Industries. On the one hand, the Internet of Things, artificial intelligence and other technologies penetrate into high energy-consuming industries, and reduce the intensity of energy consumption through intelligent transformation. For example, rural digitalization increases agricultural carbon productivity by 0.629% through scale operation[3]; On the other hand, the digital industry's own low-carbon characteristics promote the industrial structure to lean towards low-carbon. The non-resource-based cities have a more flexible industrial structure, and the digital economy has a significantly higher promotion effect on carbon productivity than the resource-based cities[8]. Based on this, hypothesis 3 is proposed.

Hypothesis 3: Optimization of industrial structure is the transmission channel for digital economy to affect carbon productivity.

4. Research Design

4.1 Model Design

4.1.1 Baseline regression model

This paper constructs a double fixed effect model for empirical test, and uses the previous practice of domestic scholars for reference to set up the following benchmark linear regression model:

$$Yit=\alpha_0+\alpha_1 Xit+Controls_{it}\alpha_2+\gamma_i+\theta_t+\xi_{it} \quad (1)$$

In the formula (1), Yit is an explanatory variable, Xit is an explanatory variable and Cit is a control variable; Cityi and yeart represent fixed effects of city and year respectively, and ξit represents random disturbance term.

4.2 Variable Description

4.2.1 Dependent variable

Carbon productivity (lnCP). Referring to the research of Kaya et al. (1997), the natural logarithm of GDP output level per unit of carbon emission is used as the proxy indicator of carbon productivity, and the carbon dioxide emissions of each city are calculated by the emission factor method published by the United Nations Intergovernmental Panel on Climate Change (IPCC). The specific formula is:

$$lnCP=ln(GDP/CDE) \quad (2)$$

In formula (2), GDP is the actual GDP of each region; CDE is carbon dioxide emissions. In this paper, the method in the 2006 IPCC Guidelines for National Greenhouse Gas Inventories is used to calculate the carbon dioxide emissions, which are converted from the consumption of 8 major fossil energy sources.

4.2.2 Explanatory variables

Digital economy (digi). At present, there is no final conclusion on the calculation of the digital economy. This paper refers to Zhao Tao et al.[20] (2020) and selects five indicators, namely, the number of Internet users per 100 people, the proportion of computer service and software employees, the total number of telecommunication services per capita, the number of mobile phone users per 100 people and the Peking University Digital inclusive finance Index, and finally obtains the digital economy index by entropy method, as shown in Table 1.

Table 1 Evaluation System of Digital Economy Development Level

Level 1 indicators	Secondary indicators	Indicator description	Indicator attribute
Internet development	Internet penetration	Internet users per 100 population	+
	Number of Internet-related employees	Percentage of Computer Service and Software Professionals	+
	Internet-related outputs	Total telecom business per capita	+
	mobile Internet	Number of mobile phone users per 100 population	+
digital financial inclusion	Digital inclusive finance Development	Peking University Digital inclusive finance Index	+

4.3 Control Variables and Intermediate Variables

Education level (lnedu) is controlled by logarithm of education expenditure, government environmental regulation (lnErs) is controlled by logarithm of sewage treatment rate, research and development investment (lnrd) is controlled by logarithm of number of R&D personnel, industrial structure (lnis) is controlled by logarithm of total value of secondary industry as a proportion of GDP, and population density (pop) is controlled by ratio of number of people to area. In this paper, technology innovation (lninno) is selected as an intermediate variable to test the mechanism, and the number of patents granted in each region is added to 1 for logarithmic processing.

4.4 Data Sources

This paper selects the panel data of 274 prefecture-level cities in China from 2011 to 2022, with a total of 3288 observations.

The digital inclusive finance data are derived from the Digital Finance Research Center of Peking University, and all other original data are derived from China Statistical Yearbook, Local Bureau of Statistics and statistical yearbooks of various regions. Some missing values are supplemented by linear interpolation.

5. Empirical Results Analysis

5.1 Descriptive Statistics

The descriptive statistics of the variables are shown in Table 2. As can be seen from Table 2, the average value of carbon productivity in all samples is 0.0969, the maximum value is 0.4581, and the minimum value is only 0.006. The standard deviation represents the degree of dispersion of variables. The maximum standard deviation in the above table is 9.9059, which reflects the uneven development of digital economy among regions and may have heterogeneous impact on carbon productivity.

Table 2 Variables and Descriptive Statistics

Variable type	variable	Number of observations	average/mean value	standard deviation	minimum value	maximum
Explanatory variable	digi	3288	0.0969	0.0812	0.006	0.4581
Interpreted variable	lnCP	3288	8.2644	0.8499	6.3559	10.4243
	lnis	3288	3.7741	0.2756	2.4596	4.4073
	lnedu	3288	13.2143	0.7868	9.9059	16.2761
Control variable	lnErs	3288	4.475	0.1779	2.8903	4.804
	lnrd	3288	8.7753	1.6887	2.5649	13.2609
	pop	3288	5.7681	0.9137	1.6094	7.9423
mediator variable	lninno	3288	5.3672	1.8934	0	11.3866

The high difference between research and development investment and education level (lnedu) in the control variables may lead to collinearity, which needs to pass the VIF test. Therefore, the VIF test is used to perform multicollinearity test on all variables before regression. The results are shown in Table 3, with VIF of all variables less than 10, indicating that there is no multicollinearity when the test passes.

Table 3 Multivariate Collinearity Test

variable	VIF	1/VIF
pop	1.610	0.620
digi	1.210	0.829
lnErs	1.120	0.890
lnis	1.080	0.926
lnedu	1.060	0.944
MeanVIF		1.330

5.2 Baseline Regression Results

This paper selects two-way fixed effect model to test the impact of digital economy on urban carbon productivity. The specific results are shown in Table 4. The above regression results show that the digital economy has a significant positive impact on carbon productivity, and with the gradual addition of control variables, the coefficient and significance of the digital economy show a stable and increasing trend. Columns (1) to (6) of Table 4 are the regression results after adding the control

variables in turn. The coefficient of Digital Economy (digi) is still significantly positive, and adding all the control variables is significant at the 5% level. This indicates that the digital economy can significantly promote carbon productivity, which verifies Hypothesis 1.

Table 4 Benchmark Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)
digi	**0.434***	**0.433***	**0.537****	**0.542****	**0.560****	**0.557****
	(1.80)	**(1.79)**	**(2.25)**	**(2.27)**	**(2.36)**	**(2.34)**
lnedu		0.0343	0.00785	0.00844	0.00615	0.00536
		(0.84)	(0.19)	(0.21)	(0.15)	(0.13)
lnis			0.439***	0.438***	0.436***	0.433***
			(7.62)	(7.61)	(7.54)	(7.37)
lnErs				-0.0183	-0.0177	-0.0181
				(-0.28)	(-0.27)	(-0.27)
pop					0.162	0.160
					(0.97)	(0.97)
lnrd						0.00669
						(0.49)
_cons	8.222***	7.769***	6.451***	6.528***	5.631***	5.606***
	(331.21)	(14.43)	(11.82)	(10.28)	(4.82)	(4.80)
City	control	control	control	control	control	control
year	control	control	control	control	control	control
N	3288	3288	3288	3288	3288	3288
r2	0.908	0.908	0.910	0.910	0.910	0.910
r2_a	0.899	0.899	0.902	0.902	0.902	0.902

T statistics in parents, * p < 0.1, * * p < 0.05, * * * p < 0.01, the following table is the same.

5.3 Heterogeneity Analysis

Due to the significant heterogeneity in resource endowments, location conditions and scale levels of each city in China, this heterogeneity is not only the objective basis for differentiation of urban development, but also shapes the uniqueness of different cities in industrial layout, growth model and transformation path. Therefore, we examine the impact of digital economy on carbon productivity under the heterogeneous resource and city scale. According to the regional planning standard of China, the city is divided into two regions: resource-based city and non-resource city. Municipalities directly under the central government, provincial capital cities and sub-provincial cities are divided into central cities and the rest are peripheral cities.

5.3.1 Analysis of resource heterogeneity

Table 5(1)(2) shows the heterogeneous results of resource-based cities and non-resource-based cities respectively. It is found that the digital economy only has a positive significant impact on non-resource-based cities, while the impact on resource-based cities is not significant. This is mainly because the economy of resource-based cities is highly dependent on a single natural resource, such as mineral resources and energy, with a single industrial structure, difficult industrial transformation and upgrading, backward construction of digital infrastructure and poor circulation of data, capital and talents, which will restrict the promotion of the digital economy on the carbon productivity of resource-based cities. Rather than a resource

city, it will gather more universities, scientific research institutions and some innovative enterprises, have better digital infrastructure, can quickly realize the digital transformation of manufacturing industry, improve production efficiency and energy use efficiency, and can better meet the needs of the development of digital economy, so that the digital economy has a significant positive effect on its carbon productivity.

5.3.2 Analysis of resource heterogeneity

From (3)(4), it can be seen that the digital economy has a positive impact on the carbon productivity of peripheral cities, but this effect is not significant for central cities. This may be due to the low operating cost of constructing data centers in peripheral cities, which attracts the business transfer of digital economy enterprises to realize a significant decrease in carbon emissions and an increase in economic benefits at a lower cost. However, central cities often face more land and labor costs, and it is difficult to reduce carbon and increase efficiency in the digital economy because of land and energy costs. On the other hand, the industries in the central cities are more inclined to the high-tech industries, and their own carbon emissions may be low, while the industrial structure in the peripheral cities is mainly dominated by the traditional high energy consumption, thus the space for carbon reduction is larger, and the marginal utility brought by the emission reduction in the digital economy is larger.

Table 5 Results of Heterogeneity Analysis

	(1)	(2)	(3)	(4)
	Resource-based city	**Non-resource city**	**key city**	**Peripheral cities**
digi	0.298	0.770**	0.762	0.522**
	(0.90)	(2.40)	(1.34)	(2.01)
lnis	0.231**	0.577***	0.721***	0.401***
	(2.55)	(7.68)	(5.49)	(6.31)
lnErs	0.0550	-0.0673	0.0653	-0.0231
	(1.03)	(-0.71)	(0.69)	(-0.33)
lnrd	0.0231	-0.00496	-0.00096	0.00344
	(0.91)	(-0.32)	(-0.04)	(0.21)
lnedu	0.109*	-0.0492	0.111	-0.0144
	(1.68)	(-0.94)	(1.00)	(-0.33)
pop	-0.319	0.352*	0.438	-0.0197
	(-1.05)	(1.76)	(1.48)	(-0.09)
_cons	7.356***	4.930***	1.197	7.054***
	(3.91)	(3.29)	(0.47)	(4.80)
City	control	control	control	control
year	control	control	control	control
N	1152	2136	396	2892
r2	0.924	0.902	0.929	0.908
r2_a	0.915	0.892	0.918	0.899

5.3 Robustness Test

In the three models, the coefficient of digi is positive, and significant at least at 5%, indicating that the positive impact of digital economy on carbon productivity is robust. Table 6(1) shows that after excluding the municipalities directly under

the central government, the coefficient slightly decreased (0.518), but the significant increase may be due to the special economic structure of the municipalities directly under the central government, such as policy inclination and concentration of resources, and the results after excluding more reflect the general laws of ordinary regions. Table 6(2) shows that after replacing the standard error, the coefficient slightly increased (0.557), but the significance decreased. It may be because the original standard error underestimated the error. After replacing the robust standard error, the value of T was close to the critical value and remained marginal significance. Table 6(3) After adjusting the sample period, the coefficient increased significantly (1.049), which should be due to the fact that specific events in the sample period strengthened the effect of digitalization, or excluded the data of interference period, highlighting the core effect.

Table 6 Results of Robustness Test

	(1)	(2)	(3)
	Excluding municipalities directly under the central government	**Replacement standard error**	**Adjust sample period**
digi	0.518**	0.557*	1.049**
	(2.17)	(1.96)	(2.54)
Control	control	control	control
City	control	control	control
year	control	control	control
N	3240	3288	2740
r2	0.910	0.910	0.900
r2_a	0.901	0.902	0.888

5.4 Mechanism Analysis

Table 7, column (1), reports the results of the mechanism test of technological innovation as an intermediate variable. The results show that the positive correlation between digital economy and technological innovation is significant. The coefficient of digital economy (digi) is 0.139, which is significant at the statistical level of 1%, indicating that for every unit of increase in the level of development of digital economy, the level of technological innovation will increase by 0.139 units on average. This result directly verifies the promotion effect of digital economy on technological innovation, which is in line with the logic of "digital economy promotes research and application of green technology through data element integration and interdisciplinary innovation ecological construction" in the theoretical mechanism. After adding the control variables, the research and development investment (lnrd) is significantly positive, indicating that research and development investment is an important driver of technological innovation and forms a synergistic effect with the digital economy; Population density (pop) is significantly negative, which may reflect the crowding-out effect of overpopulation on innovation resources. Generally, hypothesis 2 can be proved to be true.

Table 7 (2) reports the results of the mechanism test of industrial structure upgrading as an intermediate variable. The results show that the positive correlation between digital economy and industrial structure upgrading is significant. The coefficient of Digital Economy (digi) is 0.00238, which is significant at 1% of the statistical level, indicating that the development of digital economy can promote the transformation of industrial structure to low-carbon and service-oriented. This is consistent with the expectation of "digital technology permeates traditional industries and promotes industrial synergy" in the theoretical mechanism. After adding the control variables, the industrial structure (lnis) is significantly negative at 10%, which may reflect that the proportion of high-energy consuming industries in the current industrial structure is still high, and it has certain inhibitory effect on upgrading, but the positive effect of digital economy can partially offset this effect. Generally, hypothesis 2 can be proved to be true.

Table 7 Mechanism Test Results

	(1)	(2)
	technical innovation	**upgrading of an industrial structure**
dig	0.139***	0.00238***
	(4.79)	(4.88)
lnis	-0.337	-0.0111*
	(-0.62)	(-1.72)
lnErs	0.00849	0.000347
	(0.26)	(0.62)
lnrd	1.043***	0.00413
	(2.90)	(0.69)
lnedu	0.624	0.00662
	(1.47)	(1.50)
pop	-3.99***	-0.00927
	(-3.89)	(-0.62)
_cons	5.878	0.0511
	(1.44)	(1.16)
City	control	control
Year	control	control
N	3288	3288
r2	0.973	0.635
r2_a	0.968	0.573

6. Conclusions and Implications Recommendations

6.1 Conclusion

The digital economy has significantly promoted the growth of total factor carbon productivity through green technological innovation and industrial structure optimization. For example, the National Big Data Comprehensive Test Area policy has effectively reduced the growth rate of urban carbon emissions and increased the growth rate of carbon productivity by promoting green technological innovation, extensive application of robots and optimization of industrial structure[5]. The impact of the development of the Internet on carbon productivity has a double-threshold effect. In the mid-term of industrialization, the carbon productivity is increased through resource allocation efficiency, and the effect is further strengthened through technological progress in the later stage. Rural digitalization promotes agricultural total factor carbon productivity through technological innovation, scale operation and non-agricultural employment, and there is a significant spatial spillover effect.[13]. In terms of regional differences, the promotion of digital economy is more significant in the eastern region and regions with higher levels of new infrastructure. For example, the enhancement effect of carbon productivity is more prominent in the eastern region due to the high level of collaborative innovation and the strong ability of factor agglomeration in the digital industry; In regions with high levels of new infrastructure, the integration of digital technology and industry is more efficient, and the efficiency of resource allocation has been significantly improved. In the middle stage of industrialization, the digital economy mainly improves carbon productivity by optimizing the allocation of resources; In the latter stage of industrialization, technological progress has become the main driving force[10]. Platform economy has a “masking effect” in the impact of digital construction on carbon productivity, but after the implementation of the “double carbon” strategy, the effect is weakened and the positive effect of digital is significantly enhanced.

6.2 Revelation suggestion

First, at the government level, we must first improve the layout of the digital infrastructure. Increase investment in digital

infrastructure in the central and western regions to narrow the regional "digital divide." For example, reference is made to the experience of building a green data center in the National Big Data Comprehensive Test Area[5], to lay out low-carbon data centers in the central and western regions, promote the application of renewable energy and reduce the energy consumption of data centers. At the same time, we will speed up the construction of 5G, industrial Internet and other new infrastructure to support the industry's digital transformation[11]. Thirdly, differentiated policy guidance and regional coordination. For the eastern region, we will focus on supporting the construction of digital industry collaborative innovation clusters, and give full play to its advantages of technology spillovers and factor agglomeration[7]; For the central and western regions, relying on the opportunity to undertake industrial transfer, we will promote the digital transformation of traditional industries, such as using big data to optimize the agricultural production process[20]. It is also necessary to establish a trans-regional digital economic cooperation mechanism. For example, regions such as the Yangtze River Economic Belt can amplify the spatial spillover effect through data sharing and technology collaboration. It is also very important to strengthen policy incentives and environmental regulations. The government has formulated carbon productivity assessment indicators, incorporated the digital economy into the green development policy system, and granted tax relief and financial subsidies to enterprises in their digital transformation[11]. With the help of the "two-carbon" strategy, the platform economy will be forced to become green, the expansion of its high-energy consumption business will be restrained, and the flow of resources to low-carbon industries will be guided.

Second, at the social level, we must first promote collaborative innovation in Industry-University-Research. Encourage universities, scientific research institutions and enterprises to cooperate, establish a digital technology innovation alliance, and focus on tackling low-carbon technologies. Such as intelligent energy management and green manufacturing technology. Thirdly, it is to enhance public awareness of low carbon and digital literacy. Through popular science publicity and education guidance, enhance the awareness of enterprises and the public on the low-carbon value of the digital economy. For example, the "green consumption" model in internet plus was promoted, e-commerce platforms were encouraged to carry out green product traceability, and consumers were guided to choose low-carbon goods[19]. The society also needs to perfect the financial support system. Through the development of digital inclusive finance, the financing constraints of the digital transformation of small and medium-sized enterprises will be eased. For rural areas, scale operations and green technology applications can be supported through digital financial instruments.

Third, at the enterprise level, it is first necessary to promote the application of production intelligence and green technology[21]. Manufacturing enterprises should introduce industrial robots, internet of things and other technologies to optimize production processes to reduce energy consumption. For example, learn from the experience of robot application in the big data test area To reduce carbon emissions from highly polluting sectors through automated production. Agricultural enterprises can rely on the rural digital platform to achieve accurate fertilization and intelligent irrigation, and improve the efficiency of resource utilization[13]. Thirdly, optimize the industrial structure and coordinate agglomeration. Traditional enterprises are transforming to service-oriented manufacturing and developing digital service industries with high added value and low energy consumption, such as industrial internet platforms. Participate in industrial synergy, such as joining the "manufacturing-producer services" collaboration network, and reduce overall carbon emissions through division of labor and technology sharing[1]. But also to strengthen data-driven carbon management capabilities. Establish an enterprise carbon footprint monitoring system and optimize the energy consumption structure by using big data analysis. For example, logistics companies can reduce transportation carbon emissions through intelligent scheduling systems, and e-commerce platforms can reduce packaging waste through supply chain digitization[22].

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Zhang, L. J., & Zhang, Y. L. (2025). Digital economy, high-tech industrial agglomeration and carbon productivity.

Statistics and Decision, 41(08), 23–28.

[2] Ma, S. J., He, Z., & Xu, Y. (2025). Digital construction, platform economy and total factor carbon productivity. Statistics and Information Forum. Retrieved May 20, 2025, from 1-13.

[3] Hou, J., & Lu, M. Y. (2025). Threshold effect analysis of carbon productivity in digital rural empowered agriculture. Statistics and Decision, 41(02), 42–47.

[4] Lin, Y., & Liu, Z. J. (2025). Research on the impact of digital economy on urban carbon productivity. Economic Problem Exploration, (01), 154–171.

[5] Chen, H. B., & Zhang, X. X. (2024). How the national big data comprehensive test area promotes urban green transformation: Based on the dual perspectives of carbon emission growth and carbon productivity growth. China Soft Science, (11), 172–188.

[6] Gong, H. Y., Ni, C., & Cao, Y. L. (2025). Digital economy and carbon productivity: Evidence from logistics industry. Environmental Science. Retrieved May 20, 2025, from 1-14.

[7] Xiong, L., Zhang, Q., Yan, Y. L., et al. (2024). Spatial and temporal characteristics and driving factors of coupling and coordination between the development level of urban agglomeration in the middle reaches of the Yangtze River and carbon productivity. Economic Geography, 44(10), 46–54.

[8] Chen, S. H., & Xu, Y. D. (2024). Can the digital economy effectively increase the provincial carbon productivity? Southeast Academic, (05), 80–92+246.

[9] Peng, W. B., & Cao, X. T. (2024). Research on the impact of digital infrastructure construction on urban carbon productivity: Empirical evidence from "Wideband China". Economic Journal, (08), 94–105.

[10] Li, H. Y., Wang, Q. Y., & Lu, F. Z. (2024). How the digital economy affects industrial carbon productivity: A biased perspective based on technological progress. Financial Science, (08), 116–134.

[11] Zeng, H. Y., & Liu, J. L. (2024). Research on the impact of China's industry digitalization on improving carbon productivity: Based on the consideration of industrial structure optimization and environmental regulation effect. Price Theory and Practice, (03), 107–111+222.

[12] Wang, D., & Liu, H. H. (2025). Digital economy empowers low carbon development in regions: Impact effects and mechanism. Journal of China University of Mining and Technology (Social Science Edition), 27(01), 158–175.

[13] Wang, F. T., Wang, H., & Kong, F. B. (2024). The effect of rural digital development on agricultural total factor carbon productivity. China Population, Resources and Environment, 34(03), 79–90.

[14] Bai, W. T., Chen, J. C., Hou, J., et al. (2025). How does digital transformation affect agricultural carbon productivity?: Empirical evidence from China. Agricultural Resources and Regionalization in China. Retrieved May 20, 2025, from 1-16.

[15] Cao, H. Y. (2024). Research on the impact of collaborative innovation of digital industry on total factor carbon productivity. Modern Management Science, (01), 74–82.

[16] Xie, F. J., & Jiang, X. J. (2024). Digital economy, industrial intelligence and industrial carbon productivity: A time-space analysis based on the Yangtze River economic belt. Modern Management Science, (01), 187–198.

[17] Wu, C. Q., & Deng, M. L. (2023). Research on the impact of digital economy development on industrial carbon productivity in China. China Soft Science, (11), 189–200.

[18] Huang, Y. N. (2023). Digital economy, low-carbon technological innovation and total factor carbon productivity. Technology Economy and Management Research, (08), 26–32.

[19] Hu, H. H., & Shen, J. (2023). Digital economy, green innovation and "Double Carbon" target: "Emission Reduction" and "Efficiency Enhancement" perspective. Journal of Nanjing University of Finance and Economics, (04), 79–88.

[20] Kaya, Y., & Yokobori, K. (1997). Environment, energy, and economy: Strategies for sustainability. United Nations University Press.

[21] Fang, F. Q., Tian, G., & Zhang, X. (2023). Digital infrastructure and intergenerational income upward mobility: A quasi-natural experiment based on the "broadband China" strategy. Economic Research, (5), 79–97.

[22] Liu, Y., Wu, Y. Q., & Zhu, Z. H. (2023). Digital economy and urban carbon performance enhancement: Effects and mechanisms. Research World, (08), 69–78.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Inheritance Path of Intangible Cultural Heritage under the Strategy of Cultural Digitalization —Taking a Number of Intangible Cultural Heritage in Shandong Province as an Example

Yaning Ren*, Xiaoteng Huang

School of Public Administration, Shandong Technology and Business University, Yantai, Shandong, 264005, China

Corresponding author: Yaning Ren

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This article takes the cultural digitalization strategy as the entry point, and takes the intangible cultural heritage in Shandong as an example to explore its contemporary inheritance and diversified practice strategies. Studies have shown that intangible cultural heritage is an important carrier of regional culture, and current challenges facing intangible cultural heritage—such as insufficient inheritance space, narrowing transmission areas, and limited communication channels—coincide with the opportunities presented by the digital era. Shandong Province has explored Lu brocade and gourd carving from the technical level, platform construction, collaboration with industries, etc., but issues such as uneven development, redundant efforts, insufficient coordination, and lack of talent remain. This article systematically identifies that intangible cultural heritage must achieve a balance among three aspects: protection, innovation, and communication. At the level of protection, focus on preserving the core content and technical details of intangible cultural heritage through standardized digital recording methods; At the level of innovation, promote the integration of skills with modern media and life scenes to enhance its contemporary vitality; At the communication level, build a hierarchical and classified communication system that connects internal and external to enhance the social cognition and cultural influence of intangible cultural heritage. This paper believes that the inheritance of intangible cultural heritage should always be based on cultural connotation and human inheritance as the core, so as to avoid falling into the misunderstanding of technology supremacy or commercialization. Shandong's practice shows that digital technology can provide important support for the survival and activation of intangible cultural heritage, but it must be combined with the law of inheritance. In the future, we should further improve the mechanism construction, promote multi-party cooperation, and promote the real living inheritance and possibility of intangible cultural heritage in contemporary society.

Keywords: Cultural Digitalization Strategy; Shandong Intangible Cultural Heritage; Intangible Cultural Heritage Protection

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1414

1. Introduction

In the face of the current wave of globalization and digitalization, as a carrier of cultural diversity and national memory, the protection and inheritance of intangible cultural heritage has ushered in more opportunities and challenges. With the

implementation of the national cultural digitalization strategy and the actual needs of intangible cultural heritage protection, the digitalization of intangible cultural heritage is a new path to surpass the "bottleneck" of protection and stimulate the "power" of inheritance to break through the barriers of intangible cultural heritage protection [1]. The 20th National Congress of the Communist Party of China proposed "implementing the national cultural digitalization strategy" for the protection of intangible cultural heritage, which has programmatic and policy significance for the digitalization of intangible cultural heritage protection, and has changed from "rescue protection" to "preventive protection and creativity" for the protection of intangible cultural heritage. "The systematic transformation is of great significance. Intangible cultural heritage is an important part of the excellent culture of the Chinese nation, with a long history and profound heritage. Shandong Province is a major province of intangible cultural heritage, with a large number of national and provincial intangible cultural heritage, profound historical accumulation, and rich regional culture. However, many intangible cultural heritage are still mainly passed on by word of mouth, and problems such as inheritance, limited transmission areas, and modern inheritance faults are still outstanding. Intangible cultural heritage digital technology has crossed the threshold of inheritance, enabling intangible cultural heritage to release its cultural value and economic value, testing the construction of Qilu's cultural power. At the same time, intangible cultural heritage is also a "going out" business card of Chinese culture in the era of globalization going to the world. Since the current external communication of intangible cultural heritage is still limited by language, culture, channels, etc., it is necessary to explore the paths and methods of external communication of intangible cultural heritage, create a multi-language and multi-form digital communication ecology of intangible cultural heritage [2], and increase the influence of Chinese culture. Therefore, sorting out the practical process of digital communication of intangible cultural heritage in Shandong Province, summarizing and problems, and finding possible optimization solutions not only has certain significance at the theoretical level, but also has important practical value for the contemporary inheritance of intangible cultural heritage. Inheritance and new ideas and new ideas in the digital age [3].

2. The Current Situation of Intangible Cultural Heritage Digital Practice in Shandong Province

Shandong is a province rich in intangible cultural heritage (ICH) resources. Guided by the national cultural digitalization strategy, various forms of ICH in Shandong have carried out extensive digital practices and explorations, gradually forming a "technology-platform-industry" integration model. This model has promoted the digital and intelligent transformation of ICH inheritance and protection in the region. In recent years, digital practices related to ICH in Shandong have penetrated fields such as weaving, carving, food, and folk art, giving rise to a number of representative initiatives that are well-positioned to empower ICH inheritance and innovation [4].

2.1 Weaving Techniques

The digital application of Lu brocade weaving techniques serves as a typical example in the field of weaving. To support the protection of Lu brocade as a national-level ICH item [5], Heze Mobile Communications Company leveraged new infrastructure in rural areas to build a comprehensive wired and wireless network. Using 5G technology, it established a live-streaming display platform titled "5G and Live Broadcasting," through which artisans demonstrate the entire production and manufacturing process of Lu brocade. Online sales are supported via live streaming, and a "5G Smart Workshop" has been customized and transformed to enable real-time production line monitoring and management. This initiative provides a platform to promote the healthy development of Lu brocade's cultural and creative industry, forming a complete chain of "technological support – display and communication – industrial upgrading" and realizing a closed loop that integrates technology, Lu brocade display, and industry.

2.2 Carving Techniques

Carving techniques are exemplified by Dongchang gourd carving, which has developed a digital model combining online and offline efforts. Dongchangfu District has built the country's first internet platform for the gourd cultural industry. Offline, it has integrated resources into the Chinese Gourd Cultural Park, establishing gourd IoT planting demonstration parks, digital experience centers, and other experiential projects. Online, centering on four major platforms—the whole-industry-lifecycle platform, industrial resource platform, customized service platform, online sales platform, and cultural tourism digital

platform—it has established a big data center for the gourd industry. These efforts have contributed to the development of standardized, high-quality, and brand-oriented Dongchang gourd carving techniques, creating a strong linkage between ICH and industry [6].

2.3 Folk Art

Digital practices in the field of folk art are commonly manifested through new media communication, as exemplified by inheritors of dough figurine art [7]. By opening accounts on platforms such as Douyin and Kuaishou, these inheritors upload a variety of images, texts, and video works, including demonstrations of the creation and production process of dough figurines. This approach enables low-cost dissemination of dough figurine art to a broad audience.

2.4 Regional Integrated Advancement

At the regional level, systematic digital promotion has achieved notable results. Qihe County has completed the construction of the "One Book, One Room, One Hall, One Library" digital cultural infrastructure project and has recorded 94 ICH items and 72 inheritor records on the provincial ICH digital management platform. The county also relies on WeChat official accounts, apps, and other platforms to host activities such as cloud exhibitions of ICH skills and "ICH Shopping Cloud Store Exploration," enabling systematic digital management and extensive dissemination of regional ICH resources.

In addition, the launch of the public welfare digital collection Quancheng Intangible Cultural Heritage Baby represents an innovative form of digital communication for ICH. The collection features six sets of digital images covering Jinan's ICH items, such as Lu embroidery and Zhangqiu iron pot. This allows the public to access ICH in a more intuitive and convenient way, thereby enhancing its appeal and dissemination.

3. Digital Practice of Intangible Cultural Heritage in Shandong Province

Although the digital application of intangible cultural heritage (ICH) in Shandong Province has achieved certain results, several problems remain in terms of its breadth and depth, which constrain the scope and effectiveness of digital inheritance. These issues include insufficient technical coverage and limited deep application of digital technologies.

Digital resources tend to be concentrated on well-known ICH items with high commercial value, such as Lu brocade and Dongchang gourd carving. In contrast, most small-scale ICH projects on the verge of extinction have not yet undergone digital collection or organization due to factors such as remote locations or insufficient funding, and may gradually be forgotten. At the same time, the digitization of most projects remains at the stage of preserving original records—such as audio and video documentation—with limited engagement with emerging technologies like VR/AR, virtual reality, and digital twins. The focus remains largely on knowledge transmission and manual operation demonstrations, while digital outcomes lack cultural presentation and immersive experiences, failing to reflect the deeper cultural dimensions of ICH.

At the level of content dissemination and transformation, there are issues of homogeneous communication and weak industrial transformation chains. The digital dissemination of folk art ICH items mostly relies on general social media platforms such as Douyin and Kuaishou, with content primarily consisting of simple recordings of creative processes and finished product displays. There is a lack of in-depth interpretation of ICH cultural backgrounds, historical origins, and technical principles. The content is highly homogeneous, making it difficult to establish a differentiated communication advantage, which results in low user engagement and limited dissemination effectiveness.

Although some projects have built online display and sales platforms, they have failed to effectively integrate resources such as creative design, brand operation, and marketing. The transformation chain from digital outputs to commercial and cultural value remains incomplete. Sales are still at the preliminary stage of "live-streaming with goods," lacking higher-value transformation paths such as personalized customization or cultural and creative IP development, thus failing to fully unlock the industrial potential of ICH.

The issue of coordinating talent and resources is also prominent. On the one hand, ICH digitization requires professionals with expertise in both ICH and digital technology. Currently, Shandong Province faces a shortage of such interdisciplinary talent. ICH inheritors often lack the mindset and habits necessary for digital application and communication, and their understanding of digital culture is insufficient, leading to a lack of coordination and collaboration that hinders innovation in digital applications [8].

On the other hand, there is a lack of overall planning for digital resources across regions and ICH categories, resulting in a fragmented approach where "each operates in isolation." Although regional ICH digital management platforms have been established in areas such as Qihe County, the province as a whole has yet to develop a co-construction and sharing framework for ICH digital resources. The standards for digital outcomes vary across regions, making them incompatible with one another. High-quality resources cannot be shared or reused, undermining the advantages of digitalization. Coupled with difficulties in securing sustained funding, these challenges make it difficult to carry out comprehensive digital initiatives—including scientific research, platform development, and team building—especially for niche ICH projects, where the lack of financial support hinders the smooth implementation of digital efforts.

4. Countermeasures and Suggestions for Digitalization of Intangible Cultural Heritage in Shandong Province

For the first time, the "14th Five-Year Plan" cultural plan deploys the "intangible cultural heritage digital project" and the "national cultural big data system" side by side, which indicates that intangible cultural heritage inheritance is facing unprecedented opportunities and new challenges. As a province with a large number of intangible cultural heritage resources, Shandong Province has carried out rich intangible cultural heritage digital practices in terms of policy support, technology application, and industrial integration, providing valuable Shandong experience. For example, the Shandong Provincial Department of Culture and Tourism has established the "Shandong Intangible Cultural Heritage Digital Center", universities have set up "technical big data laboratories", and enterprises have jointly built a "Intangible Cultural Heritage Exploration Store " live streaming base in collaboration with the provincial ICH protection center, among other cooperative initiatives [9]. However, in the actual promotion process, many difficulties have been encountered, such as the imbalance between technology application and cultural core connection, the difficulty of intangible cultural heritage inheritors adapting to digitalization, the loss of value and commercialization tendency in cultural communication, and the long-term mechanism Not yet sound, etc. For example, the case of Lu brocade weaving illustrates the imbalance between technology application and cultural core, some digital projects stay in the combination and gimmicks of technology, while ignoring the ingenuity and connotation of "weft and latitude"; The old craftsmen in Lu brocade weaving do not admit that digital projects can be disseminated through various means such as live broadcasts. Young people artificially simplify the complicated weaving process in order to attract traffic, which is a little bit that needs to be improved and optimized urgently. Faced with the above-mentioned complex situation, it is particularly urgent to build a systematic, inclusive, and feasible digital inheritance path for intangible cultural heritage. This path goes beyond simple technical application thinking, based on the inherent laws of cultural inheritance, and seeks coordinated advancement in the three dimensions of protection, innovation, and dissemination.

At the level of protection, the main task is the permanent and authentic preservation of cultural heritage. To solve these problems, it is necessary to formulate a complete set of digital collection standards for intangible cultural heritage, and formulate different collection plans for different intangible cultural heritage items. Specific to the technical Zibo glass firing process, donkey-hide gelatin production, etc., various methods such as 3D scanning and inheritor oral records can be used to record tacit knowledge that cannot be fully expressed through verbal; Specific to the Qufu Confucius Festival Ceremony, Guzi Yangko etc, the cultural atmosphere and emotions of these cultural heritages can be recorded by panoramic shooting and video recording [10]. In addition, there are issues of digital ethics and intellectual property protection. It is necessary to clearly define the ownership and scope of use of intangible cultural heritage digital achievements, effectively safeguard the cultural rights and interests of inheriting groups, and avoid cultural encroachment and improper use that may arise during the digitalization process.

At the level of innovation, enrich the presentation of intangible cultural heritage, combine modern media, use the Internet and artificial intelligence, and combine intangible cultural heritage with innovation [11]. Technology brings unlimited possibilities: through VR/AR technology, the immersive experience is realized, so that the intangible cultural heritage has more possibilities, and the story behind the exhibits is linked through AR scanning, so that the audience can "substitute" the intangible cultural heritage, experience the immersive atmosphere; Use AI big data to analyze user concerns and push corresponding content. The focus of innovation is to actively integrate across borders without compromising the essence

of intangible cultural heritage, retain the rhythm and keep pace with the times, so as to make culture live and make culture "live" in modern times. Use digital language intangible cultural heritage to resonate with people's emotions, so that intangible cultural heritage can be integrated into the public in innovation.

At the level of communication, it provides vivid images, creates an atmosphere, and spreads across layers. Internally, use a series of products of the times such as social platforms to produce targeted intangible cultural heritage-themed videos. For young people: use interesting videos to spread intangible cultural heritage such as Lu brocade weaving and dough figurine making ; Provide online intangible cultural heritage teaching, online digital museums and other traceability needs for middle-aged and elderly groups. At the same time, promote the digitalization of intangible cultural heritage into the classroom, and produce digital textbooks for Shandong intangible cultural heritage cases, allowing students to experience intangible cultural heritage crafts in virtual simulation experiments; The foreign publicity language platform translates intangible cultural heritage such as Lu brocade weaving, donkey-hide gelatin production, gourd carving, and dough figurines into English, Japanese, Korean and other languages, and conducts cultural interpretation to spread cultural significance to the outside world. Using international social platforms, by telling "story crafts", with empathy and story-based expression as the soul, we will promote excellent culture to go abroad and integrate into global dialogue [12]. Protection, innovation, and communication are an organic whole that supports each other: protection lays the foundation for innovation, innovation injects vitality into communication, and communication builds consensus for protection. The rich practice of Shandong Province vividly confirms the key path for digitalization to achieve creative transformation and innovation of intangible cultural heritage. Intangible cultural heritage is the memory of history and the wisdom of the nation. Intangible cultural heritage supported by digital technology transcends the time and space constraints of inheritance space, and establishes a cultural bridge with modern times with digital filing, intelligent production, immersive experience, and embedding of the entire industrial chain, to realize the "fusion and innovation" of "memory retention".

However, the digitalization of intangible cultural heritage is not a product of simply copying technology, but a complex practice with specific cultural nature and the requirements of the times. Issues such as perfection are still urgent problems that need to be solved. It is necessary for the digitalization of intangible cultural heritage to break through the fixed thinking of its instrumental thinking and find the balance of protection, innovation, and dissemination in the inheritance of intangible cultural heritage.

5. Conclusion

The road to digital inheritance of intangible cultural heritage is long and full of room for exploration. With the advancement of digital technology, the digital inheritance of intangible cultural heritage will usher in a broader space, and technological iteration will provide tools and imagination for the digital inheritance of intangible cultural heritage. Digitalization is only a means of intangible cultural heritage inheritance, and its ultimate goal is always "people"-to enhance people's understanding of cultural heritage, to empower cultural inheritance, and to promote shared appreciation among different civilizations. Therefore, the focus of the digital inheritance of intangible cultural heritage should be on building an inclusive digital inheritance network: let intangible cultural heritage form a multiple inheritance form of "entering colleges and universities, passing on the network, and reaching users" [13]. Shandong intangible cultural heritage shows that the cultural digitalization strategy not only "retains memory", but also "releases creativity" and "rebuilds industries". Build a more scientific and complete intangible cultural heritage inheritance and communication system, put forward the "three-degree" strategy for the external communication of intangible cultural heritage, adhere to "culture-based, technology-based, and people-oriented", and let the flowers of intangible cultural heritage in the "hometown of Confucius and Mencius" bloom forever online. Intangible cultural heritage is a national historical and cultural memory and a common spiritual treasure of human civilization. In recent years, the inheritance and protection of intangible cultural heritage has always been the focus of China's attention, and a series of cultural digitalization strategic policies have been proposed. Under the background of cultural digitalization strategy, the digital era presents a new chapter for ICH inheritance. Only by adhering to the principles of creative transformation and innovation, and keeping cultural inheritance in sync with the times, can intangible cultural heritage truly "come to life " and "be passed down" in the "big data era". This is the realistic requirement of cultural inheritance, and it is also the historical

mission and cultural responsibility of our generation.

Funding

This article is from 2025 College Students' Innovation and Entrepreneurship Training Program Funded Project, Project No. X2025116880060.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Deng, Y., Li, J., Zeng, C., et al. (2026). Research evolution and framework construction of intangible cultural heritage protection and inheritance in China. Journal of Changsha University, 1–10. Retrieved January 27, 2026, from https://link.cnki.net/urlid/43.1276.G4.20260124.1240.002.

[2] Zhang, Y. (2024, October 30). Let traditional intangible cultural heritage radiate a new era style. Xinjiang Daily (Chinese Edition), p. 008. https://doi.org/10.28887/n.cnki.nxjrb.2024.004994.

[3] Li, Y. (2025). Research on the path of intangible cultural heritage creative products boosting rural revitalization in Shandong section of the Yellow River Basin. Knowledge Economy, (31), 14–16+20. https://doi.org/10.15880/j.cnki.zsjj.2025.31.001.

[4] Fan, Y. (2025, November 25). Digital technology empowers the protection and inheritance of intangible cultural heritage. Shanxi Daily, p. 011.

[5] Zhu, D. (2024). Cultural and creative brand design based on “living inheritance” of intangible cultural heritage—Taking Shandong Lu embroidery as an example. China Nationalities Museum, (23), 54–56.

[6] Cheng, L., & Guo, X. (24). Multilingual communication of “intangible cultural heritage” in the context of digitization. Culture Industry, (25), 4–6.

[7] Guo, Z. (2020). Research on the development and design of intangible cultural heritage cultural and creative products in Shandong under the background of “Internet+”. Art Education Research, (1), 114–115.

[8] Zhu, Z., Dai, L., & Zhao, X. (2024). Cultivating new-quality talents: The new mission of education in the digital age. Journal of Educational Technology, 45(1), 52–60. https://doi.org/10.13811/j.cnki.eer.2024.01.007

[9] Zhao, M. (2025). Lu brocade intangible cultural heritage inheritance enters Dezhou University classrooms: Exploring the value of traditional crafts. Textile and Apparel Weekly, (20), 6.

[10] Zhou, H., & Zhou, Q. (2024). Research on the new model of digital empowerment for the living inheritance of “Shandong Handicraft” culture. Time-Honored Brand Marketing, (7), 106–108.

[11] Xia, L., & Zhang, Y. (2025, November 27). Research on the path of AI empowering intangible cultural heritage inheritance. Chongqing Science and Technology News, p. 12.

[12] Chen, Y. (2025). Exploration of the integration of intangible cultural heritage into public spaces from the perspective of digitalization. Modern Literature and Creativity, (41), 134–136. https://doi.org/10.20024/j.cnki.CN42-1911/I.2025.41.042

[13] Xu, R. (2024, December 5). Exploring the path of cultural innovation and inheritance from the perspective of cultural power. Heilongjiang Daily, p. 8. https://doi.org/10.28348/n.cnki.nhjrb.2024.006718

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Innovative Exploration of Cross-Campus Department Collaborative Governance System in Large Grade a Tertiary Hospitals from an Integrated Perspective

Yiping Zhao*

Shanghai East Hospital (Shanghai East Hospital of Tongji University), Shanghai, 200120, China

Corresponding author: Yiping Zhao, zhaoyiping4747@sina.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Under the dual guidance of expanding high-quality medical resources and achieving balanced regional development, the "one hospital, multiple campuses" model has become a core strategy for large Grade A tertiary hospitals to break through development bottlenecks and improve service accessibility. However, problems such as physical separation of campuses, fragmented departmental governance, unbalanced resource allocation, and inadequate collaborative mechanisms have restricted the synergistic effect of "1+1>2". Based on integrated governance theory and combined with multi-campus management practices, this paper systematically analyzes the core pain points in cross-campus department collaboration, and constructs a "five-in-one" collaborative governance system covering organizational restructuring, process reengineering, digital and intelligent empowerment, performance closed-loop management, and cultural integration. The feasibility and effectiveness of the system are verified through typical hospital cases. This study provides theoretical support and practical pathways for large Grade A tertiary hospitals to resolve cross-campus coordination dilemmas and promote high-quality development.

Keywords: Integrated Governance; Large Grade a Tertiary Hospitals; Cross-Campus; Department Collaboration; Governance System Innovation

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1418

1. Introduction

1.1 Research Background

With the deepening of healthcare reform and the implementation of the Healthy China strategy, large Grade A tertiary hospitals are facing the dual challenges of excessive demand for high-quality resources and uneven regional distribution. The multi-campus model has become a mainstream trend to expand high-quality medical services, relieve pressure on core campuses, and promote the downward transfer of premium resources. By 2025, more than 70% of the top 100 Grade A tertiary hospitals in China had adopted a multi-campus layout, with some operating three to five branch campuses (Dai et al., 2018).

Nevertheless, multi-campus development is not merely a physical expansion, but a deep integration of management, resources, processes, and culture. In practice, problems including poor cross-campus coordination, ambiguous authority and responsibility, and information silos have become prominent bottlenecks restricting hospital efficiency, service quality, and

sustainable development (Yuan et al., 2023).

1.2 Research Significance

1.2.1 Theoretical Significance

Existing studies mostly focus on single-campus management or medical consortium cooperation, while systematic research on cross-campus departmental collaboration remains insufficient. From the perspective of integrated governance, this paper constructs a comprehensive cross-campus collaborative governance system, enriches the theoretical framework of multi-campus public hospital governance, and provides a reference for subsequent academic research.

1.2.2 Practical Significance

By targeting realistic pain points, this paper proposes actionable and replicable paths to optimize cross-campus collaboration, improve resource allocation efficiency, promote homogeneous medical quality, and enhance patient experience. It supports similar hospitals in breaking campus barriers and achieving integrated high-quality development.

1.3 Literature Review

1.3.1 International Research

Foreign multi-campus hospital systems are relatively mature, emphasizing unified decision-making, decentralized operation, standardized processes, and information-based collaboration. Institutions such as Mayo Clinic and Johns Hopkins Medicine have formed mature models of centralized governance and differentiated development (Schmidt et al., 2025). However, due to differences in medical systems, financing mechanisms, and regulatory environments, their experience cannot be directly transplanted into China.

1.3.2 Domestic Research

Domestic studies have explored homogenized management, flattened organizational structures, and digital empowerment in multi-campus operations. Many hospitals, including West China Hospital, Wuhan Union Hospital, and Ruijin Hospital Shanghai, have carried out valuable practices in vertical management, resource sharing, and cross-campus appointment systems. However, most studies remain fragmented and lack systematic construction of a synergistic governance system covering organization, process, technology, performance, and culture (Zhang et al., 2024).

1.4 Research Methods and Innovations

1.4.1 Research Methods

Literature research: to sort out theories of integrated governance, collaborative governance, and process reengineering.

Case analysis: to summarize practical experience from representative multi-campus Grade A tertiary hospitals.

Systematic analysis: to construct an integrated multi-dimensional collaborative governance system.

Interview research: to verify the practicality of the system through interviews with hospital managers.

1.4.2 Innovations

Perspective innovation: adopting integrated governance to break through single-dimensional analysis.

Structural innovation: proposing a five-in-one collaborative governance system.

Practical innovation: combining standardized systems with operable implementation paths.

2. Core Concepts and Theoretical Basis

2.1 Core Concepts

2.1.1 Integrated Governance

Integrated governance emphasizes unified planning, resource integration, authority clarification, and process coordination to eliminate fragmentation and form a synergistic overall effect. In multi-campus hospitals, it means unified leadership, integrated management, resource sharing, and cultural integration across campuses.

2.1.2 Cross-Campus Department Collaboration

Cross-campus department collaboration refers to the coordinated behavior of clinical and administrative departments across different campuses to achieve unified medical quality, efficient resource allocation, smooth patient flow, and standardized management.

2.2 Theoretical Basis

Integrated governance theory

Collaborative governance theory

Business process reengineering theory

Resource-based view (RBV)

3. Current Situation and Core Pain Points of Cross-Campus Collaboration

3.1 Overall Status

Although many hospitals have established unified leadership and preliminary information sharing, cross-campus collaboration generally remains at a primary stage, with insufficient depth and low efficiency.

3.2 Major Pain Points

3.2.1 Organizational Fragmentation

Unclear authority between main and branch campuses; overlapping or missing management responsibilities; redundant administrative institutions.

3.2.2 Unbalanced Resource Allocation

Uneven distribution of personnel, equipment, and technologies; low resource sharing; idle resources coexisting with shortages.

3.2.3 Inefficient and Broken Processes

Long approval cycles; repeated inspections; poor continuity of cross-campus diagnosis and treatment.

3.2.4 Information Silos

Inconsistent information systems; non-unified data standards; difficulty in mutual recognition of results.

3.2.5 Inadequate Incentives and Culture

Lack of collaborative performance assessment; weak cultural identity; low willingness to cooperate.

4. Construction of the Integrated Cross-Campus Collaborative Governance System

4.1 Organizational Restructuring: Clarify Authority and Flatten Management

Establish a multi-campus management committee for top-level decision-making.

Implement vertical line management and a dual-leadership system.

Set up comprehensive coordination offices as cross-campus hubs.

Adopt a “chief director + campus director” dual responsibility system for clinical departments.

4.2 Process Reengineering: Streamline and Integrate Full Processes

Simplify cross-campus approval with online time-limited processing.

Build a unified patient journey including appointment, diagnosis, treatment, and follow-up.

Establish an emergency collaboration mechanism for public health incidents.

4.3 Digital and Intelligent Empowerment: Break Information Barriers

Build a unified multi-campus information platform.

Formulate consistent data standards and promote mutual recognition of examinations.

Develop intelligent systems for resource scheduling, telemedicine, and quality control.

4.4 Performance Closed-Loop: Incentivize Collaboration

Incorporate cross-campus collaboration into performance appraisal.

Establish special incentives and rotation mechanisms.

Form a supervision-feedback-optimization closed loop.

4.5 Cultural Integration: Enhance Identity and Cohesion

Unify core values and service culture.

Implement unified service standards and quality norms.

Promote cross-campus communication and cultivate a collaborative culture.

5. Case Verification: Ruijin Hospital Shanghai

5.1 Hospital Overview

Ruijin Hospital operates six campuses and has achieved remarkable results in integrated governance.

5.2 Practical Measures

Restructured organizational system with clear authority lists.

Reengineered cross-campus processes and shortened waiting time.

Built a unified information platform with high mutual recognition rate.

Improved performance and cultural integration mechanisms.

5.3 Implementation Effects

Explosive growth in scientific research projects.

High homogeneity of medical quality.

Significant improvement in patient satisfaction.

Effective release of cross-campus synergy.

The case fully proves that the five-in-one collaborative governance system is scientific, feasible, and effective.

6. Implementation Safeguards and Promotion Suggestions

6.1 Safeguard Mechanisms

Institutional guarantee: improve rules and authority lists.

Talent guarantee: train professional multi-campus managers.

Funding guarantee: support informatization and equipment upgrading.

Technical guarantee: maintain stable information systems.

6.2 Promotion Strategies

Promote by classification according to hospital scale and development stage.

Carry out pilot practices first and expand gradually.

Conduct dynamic evaluation and continuous optimization.

7. Conclusion and Prospect

7.1 Conclusion

Against the background of rapid multi-campus development of large Grade A tertiary hospitals, cross-campus department collaboration has become a key issue affecting governance efficiency and medical service quality. Based on integrated governance, this paper systematically analyzes the organizational, resource, process, technological, performance, and cultural dilemmas in cross-campus collaboration, and constructs a five-in-one collaborative governance system featuring organizational restructuring, process reengineering, digital empowerment, performance management, and cultural integration. The system helps break campus barriers, realize resource integration, improve service efficiency, and promote homogeneous high-quality development. Verified by practical cases, it has strong theoretical value and practical guidance for multi-campus governance reform in China's public hospitals.

7.2 Prospect

In the future, with the further development of smart hospitals, regional medical consortia, and hierarchical diagnosis and treatment systems, cross-campus collaboration will move toward deeper integration, intelligence, and refinement.

Subsequent research can further explore the application of big data, artificial intelligence, and digital twins in cross-campus collaborative governance, establish a more refined evaluation index system, and provide more scientific decision support for the modernization of hospital governance capacity.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Dai, L., Wang, H., & Chen, J. (2018). Multi-campus hospital models: A solution to equitable healthcare access in populous regions of China. Journal of Healthcare Management, 63(4), 289-302.

[2] Schmidt, H. H. W., Brück, W., & Zimmermann, W. H. (2025). Governance Models for Academic Health Care Centers: Collaboration and Innovation in Multi-Campus Settings. Leadership Perspectives, 12(1), 45-62.

[3] Yuan, Y., Zhang, Y. Q., Zhang, M. P., & Li, J. (2023). Do hospitals need to establish multiple hospital districts? A hospital-based perspective on the benefits of scale. Frontiers in Public Health, 11, 1019331.

[4] Zhang, Y., Li, J., & Wang, Z. (2024). Digital Empowerment and Cross-Campus Collaboration in Chinese Multi-Campus Hospitals: Practices and Challenges. SCIRP Journal of Healthcare Sciences, 12(3), 78-95.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Blended Learning Model for the Course of Bridge Engineering Construction Cost Planning and Management in the Context of Big Data ——Using Plateau Complex Bridge Engineering as an Example

Xiaoyan Liu*, Yinheng Luo, Jinrong Yuan, Xiaolan Wang, Jiangjie Qin, Yao Li

School of Civil Engineering, Chengdu Technological University, Chengdu, Sichuan, 611730, China

Corresponding author: Xiaoyan Liu

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: In response to the new requirements of Emerging Engineering Education for engineering management talents—namely, "understanding technology, mastering data, and being capable of management"—traditional cost management courses face issues such as a disconnect between theory and practice, as well as weak data literacy. Taking the cost control of constructing a high-pier, long-span steel truss girder bridge in a plateau region as a case study, this paper constructs a blended teaching model characterized by "online theoretical foundation, offline project-based practice, and big data thinking throughout." The reform pathway is elaborated from four aspects: curriculum content restructuring, teaching resource development, teaching implementation process, and evaluation system reform. Research indicates that the blended teaching model based on the OBE concept significantly enhances students' autonomous learning abilities, teamwork skills, and problem-solving capabilities. Moreover, the integrated teaching of BIM technology and cost estimation helps students master core cost management skills through "learning by doing" [1-2]. Practice demonstrates that this model effectively strengthens students' cost planning awareness and their ability to solve complex engineering problems, thereby providing a reference for curriculum reform in engineering management majors.

Keywords: Big Data; Cost Planning and Management; Blended Teaching; Plateau Bridge

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1412

1. Introduction

"The world looks to China for bridges, and China looks to Guizhou." Amidst the towering mountains of Southwest China, numerous super bridges span natural chasms, showcasing not only China's technological strength in bridge construction but also raising higher demands for the cultivation of engineering management talents. The construction of high-pier, long-span arch-stiffened continuous steel truss bridges in plateau regions faces multiple challenges, including harsh natural environments, complex geological conditions, and difficulties in material transportation. The cost planning and control of such projects are highly dynamic and uncertain, creating an urgent need to cultivate interdisciplinary talents equipped with big-data thinking and whole-process cost management capabilities.

However, current university instruction in the course Construction Cost Planning and Management still faces three prominent

issues: first, the teaching content overemphasizes theoretical budgeting and lacks training in the analysis of dynamic cost data during construction; second, the teaching cases are relatively outdated and fail to capture the complexities of cost control in plateau environments; third, the instructional approach remains predominantly lecture-based, leaving students with little hands-on experience in handling real-world engineering data. Therefore, transforming real-world engineering challenges into living teaching materials has become a critical task in curriculum reform.

In recent years, universities in China have actively explored reforms in engineering cost management courses. Chongqing Jiaotong University, guided by the "Two Roads" spirit, has developed a multi-module curriculum system integrating ideological and political education, foundational knowledge, bridge and tunnel engineering, and intelligent technologies. Guizhou Vocational College of Transportation has pioneered a "one bridge, one course, one team" model, developing specialized courses centered on typical bridge types. Zhanjiang University of Science and Technology, grounded in the OBE concept, has restructured the Engineering Cost Planning and Control course, employing case discussions and diversified assessment mechanisms to encourage students to take ownership of their learning [3-4]. Beijing City University has built a "whole-process cost management course cluster" based on real projects, aiming to effectively align graduates with the demands of whole-process cost management roles [5]. These practices offer valuable insights for this study. Building on this foundation, this paper focuses on curriculum reform in cost planning and management within the context of big data, aiming to explore new pathways for blended teaching models.

2. Current Teaching Status and Problem Analysis

2.1 Teaching Content Lags Behind Industry Development

At present, cost management courses in most universities still rely heavily on traditional textbooks, with teaching content focusing primarily on quota application and cost calculation, while lacking coverage of cutting-edge areas such as BIM 5D technology, big data cost analysis, and dynamic cost control [6-7]. According to the Undergraduate Professional Guidelines for Engineering Cost in Higher Education Institutions, although no specialized course modules have been specifically added, the integration of information technology with professional teaching has been achieved through laboratory open fund projects and digital project management courses. The application of BIM technology in engineering practice aligns well with this knowledge domain [8-9]. However, in the process of plateau bridge construction, cost variables such as reduced mechanical efficiency and labor idleness caused by high altitudes, large temperature differences, and complex wind field environments are rarely addressed in existing textbooks.

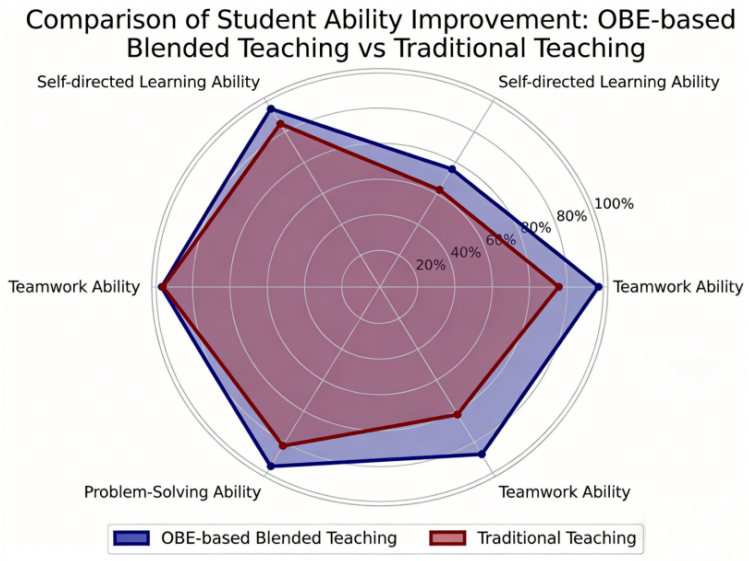
2.2 Teaching Cases Are Disconnected from Engineering Practice

The cases used in course instruction are mostly simplified exercises with limited data scales and single variable settings, making it difficult to accurately simulate the complexity of real engineering projects. Taking plateau steel truss bridges as an example, the cost planning for such structures involves multiple factors, including steel price adjustments, quota revisions for plateau regions, and the impact of special climatic conditions—all of which require analytical training based on massive amounts of real data. Currently, however, students generally lack access to real-world engineering data. Relevant research indicates that although students hold differing views on the role of building information modeling in cost overrun management, most agree that it is significantly effective in the early stages of projects.

2.3 Insufficient Cultivation of Students' Data Literacy

In the era of big data, cost managers are expected to possess comprehensive capabilities in data collection, processing, analysis, and visualization. However, current courses emphasize manual quantity takeoff and software operation while neglecting the cultivation of skills in deep data mining and pattern discovery. Students may "know how to calculate" but are not proficient in "analyzing data"; they may "understand software" but fail to grasp "the engineering logic behind the data." Research by Liu Fang et al. [1] confirms that blended teaching based on the OBE concept can significantly enhance students' autonomous learning abilities, as shown in Figure 1, providing strong support for curriculum reform.

Figure 1 Teaching Mode Competence Comparison Radar Chart



3. Construction of the Blended Teaching Model

To address the aforementioned issues, this paper takes a high-pier, long-span steel truss bridge (arch-stiffened continuous steel truss structure) in a plateau region as a running case study and constructs a blended teaching model integrating "online + offline + big data" as a trinity.

3.1 Teaching Design Framework

This model follows the design framework of "one main thread, two spaces, and three integrations":

One Main Thread: A real-world case of cost planning and control for complex plateau bridge construction runs through the entire course, forming a complete project thread from construction scheme comparison and resource allocation optimization to dynamic cost monitoring.

Two Spaces: The online space supports theoretical learning and data perception functions, including micro-lecture videos, virtual simulation experiments, and a big data case repository. The offline space facilitates case discussions and project-based practice, enabling knowledge internalization and skill enhancement.

Three Integrations: Integration of theory and practice, integration of technology and economics, and integration of online and offline learning, reflecting the organic connection among the education chain, talent chain, and industry chain.

3.2 Design of Online Teaching Modules

1. Virtual Simulation Experiment. A "virtual simulation platform for dynamic cost control of plateau bridge construction" is developed to simulate the impact of unexpected events—such as sudden climate changes, material supply delays, and equipment failures—on students' cost control decisions. Students act as cost engineers in the virtual environment, developing adjustment plans in response to emergencies. The system provides real-time feedback on cost changes, enabling students to "preview construction schemes in a virtual space."

2. Big Data Case Repository Development. Based on actual plateau bridge engineering projects, cost data generated during construction (processed to ensure confidentiality) are collected and organized to establish a "big data case repository of bridge construction costs." This repository contains multi-dimensional cost data, including labor, materials, machinery, preliminary expenses, and management costs, across different construction stages, altitude conditions, and bridge structural types, providing support for students'data analysis training.

3.3 Design of Offline Teaching Modules

1. Flipped Classroom. Before class, students engage in online learning of relevant theoretical knowledge and complete preliminary data analysis. Class time is then dedicated to discussions focused on cost planning challenges. For example, for the hoisting scheme of a plateau steel truss bridge, data on machine-shift costs of various hoisting equipment and work efficiency under different climatic conditions are provided before class. Classroom discussion centers on "how to select the

optimal machinery combination."

2. Project-Based Learning (PBL). Students are divided into groups acting as "cost consulting advisory teams," engaging in project-based practice centered on cost planning tasks throughout the entire process of plateau bridge construction. Each group is responsible for one construction phase (e.g., foundation construction, pier construction, steel truss assembly, arch-stiffening installation), completing full-process tasks such as cost budgeting, process cost monitoring, cost overrun analysis, and corrective measure formulation. Research indicates that integrated BIM project practice helps cultivate students' cost overrun management skills.

3. Dual-Tutor System Involving Both University and Enterprise. Cost managers involved in plateau bridge construction are invited to serve as industry mentors, participating in classroom discussions online or offline to share practical cost control cases and strategies from real engineering projects. Practical experience from Zhanjiang University of Science and Technology demonstrates that the involvement of industry mentors effectively enhances students' practical abilities and professional identity.

4. Teaching Implementation Process — Taking "Cost Planning for Steel Truss Girder Hoisting on a Plateau Bridge" as an Example

4.1 Pre-Class Guided Learning (Online)

The instructor assigns learning tasks: watch the micro-lecture on "Economic Comparison of Steel Truss Girder Hoisting Schemes," consult the actual cost data of hoisting construction under similar altitude conditions in the case repository, and preliminarily complete the "Cost Estimation Form for Steel Truss Girder Hoisting Scheme of a Plateau Bridge."

4.2 In-Class Internalization (Offline)

Contextual Introduction: A real case is presented — "During the hoisting of a steel truss girder for a plateau bridge, sudden strong winds caused five consecutive days of work stoppage, resulting in a cost overrun of 500,000 yuan due to project delays." Students are asked to analyze the causes of the overrun and propose countermeasures as consulting advisors.

Group Discussion: Each group uses data from the case repository and knowledge acquired online to analyze the cost composition from perspectives such as idle labor costs, equipment downtime costs, and overhead allocation. They calculate various losses using Excel or simple data analysis tools and discuss preventive measures for similar situations in the future.

Results Presentation: Each group presents their analysis report in five minutes, focusing on the main causes of cost overruns, calculation bases, and corrective recommendations.

Expert Commentary: The instructor provides feedback by referencing practical engineering solutions and introduces cases such as "how Beidou satellite-based real-time positioning technology improves hoisting efficiency," deepening students' understanding of the relationship between technology and economics.

4.3 Post-Class Extension (Online + Offline)

Each group revises and improves their cost analysis report. On the virtual simulation platform, students attempt to adjust the hoisting plan (e.g., changing machinery types, adding backup equipment) and observe the resulting cost variation curves. They then draft a "Recommendations Report on Cost Risk Prevention and Control for Plateau Bridge Construction."

5. Reconstruction of the Teaching Evaluation System

To meet the requirements of the blended teaching model, course evaluation has shifted from a "final-exam-decides-everything" approach to a whole-process, multi-dimensional assessment.

Formative Assessment (60%):

Online learning progress and quizzes (10%): assesses students' mastery of basic knowledge.

Quality of data analysis reports (30%): focuses on the accuracy of data processing and the rigor of analytical logic.

Group project outcomes and discussion performance (20%): emphasizes teamwork and problem-solving skills.

Practice from Zhanjiang University of Science and Technology [3-4] shows that using online quiz platforms to monitor students' learning progress in real time enables targeted teaching adjustments.

Summative Assessment (40%):

A major assignment titled “Cost Planning Scheme Design” based on real data is adopted. Students are required to complete a comprehensive design for a given plateau bridge construction scenario, including cost budgeting, process monitoring plans, and risk response strategies, with the focus on assessing their comprehensive application abilities.

6. Conclusion

In the context of big data, curriculum reform in bridge engineering construction cost planning and management is imperative. The blended teaching model, using complex plateau bridge engineering as a carrier—characterized by online theoretical foundation, offline project-based practice, and big data thinking throughout—can effectively address the disconnect between traditional teaching and practical application. Driven by real-world cases, students not only master the fundamental methods of cost planning but also develop their abilities to analyze and solve complex engineering problems. However, the implementation of this model places higher demands on instructors’ practical engineering capabilities and information technology literacy. How to acquire and continuously update real-world engineering big data, and how to achieve seamless integration between online and offline teaching, remain issues that require further exploration.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Liu, F. (2025). SPOC + Rain classroom based on OBE concept + Exploration of flipped classroom in teaching measurement and billing of construction projects. In Proceedings of the 2025 4th International Conference on Educational Innovation and Multimedia Technology (EIMT 2025). Atlantis Press.

[2] Wu, Y. W., Wen, M. H., Chen, C. M., & et al. (2016). An integrated BIM and cost estimating blended learning model—Acceptance differences between experts and novice. EURASIA Journal of Mathematics, Science and Technology Education, 12(5).

[3] Wan, L., & Bai, Y. (2022). Curriculum design and organization of application-oriented undergraduate engineering cost management based on BIM—Taking Zhanjiang University of Science and Technology as an example. Journal of Higher Architectural Education, 31(2), 137–144.

[4] Huang, M. C., Huang, J. G., Lan, R. S., & et al. (2022, June 8). University-level high-quality resource sharing course—Engineering cost planning and control (Engineering cost management). School of Architecture and Engineering, Zhanjiang University of Science and Technology.

[5] Zhang, W. H., Li, Z. G., Sui, Z. L., & et al. (2022). Implementation and application of whole-process cost management in the practice teaching of engineering cost major—Taking Beijing City University as an example. Theory and Practice of Innovation and Entrepreneurship, 5(18), 170–172.

[6] Jiao, L. D., Qin, M., Liu, G. F., & et al. (2019). An exploratory study on the “dual-main” teaching model in the teaching reform of engineering management majors—Taking engineering cost planning and control as an example. Modernization of Education, (66).

[7] Zhou, J. T., Zhang, H., Yang, J., & et al. (2025, April 2). Harnessing the “two roads” spirit to cultivate talents for mountainous regions—Chongqing Jiaotong University builds a talent support system for bridge and tunnel construction and maintenance in mountainous areas. China Education Daily.

[8] Wu, Y. F. (2024, March 13). “One bridge, one course, one team” assists in the construction of high bridges in mountainous areas. China Education Daily, p. 7.

[9] Luo, S., Hong, D., Du, B., & et al. (2025, November 27). The power of universities behind “bridges seeing Guizhou”. Guizhou Daily.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on Exchange Rate Risk Management in Corporate Foreign Currency Debt Financing: A Case Study of Evergrande Group

Luyao Liang*

School of International Business, Yunnan University of Finance and Economics, Kunming Yunnan, 650221, China

Corresponding author: Luyao Liang

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Amid China's ongoing reform and opening-up, the scale of foreign currency debt financing among Chinese enterprises has expanded steadily. In this context, firms face two core challenges: mitigating exchange rate risks via financial markets and managing foreign currency debt effectively. Taking Evergrande Group as a case, this paper systematically examines the exchange rate risk management framework for foreign currency debt financing and identifies key risk exposures in the process. Based on an analysis of Evergrande's current status of foreign currency debt financing, practical risk management measures, and existing deficiencies, this study proposes targeted optimization strategies. Findings indicate that strengthening the exchange rate risk management system, optimizing hedging instruments, and improving performance evaluation mechanisms have significantly reduced Evergrande's exchange rate vulnerabilities. Finally, this study extracts lessons and best practices from the Evergrande case to summarize core principles of exchange rate risk management for enterprises relying on foreign currency liabilities. It provides actionable guidance for peer firms, helps small, medium, and micro enterprises (SMEs) enhance resilience against exchange rate volatility, and supports the sound and sustainable development of foreign currency debt financing activities in China.

Keywords: Foreign Currency Debt; Exchange Rate Risk; Risk Management Strategy; Hedging; Case Study

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1409

1. Introduction

With deepening economic globalization, enterprises operate in an increasingly complex and volatile financial environment. Foreign currency debt financing has become a critical strategy for firms to diversify funding sources and lower financing costs. Nevertheless, foreign currency borrowing entails substantial exchange rate risk; large unexpected exchange rate movements can severely impair corporate financial stability. As a leading real estate developer in China, Evergrande Group relied heavily on foreign currency debt to fund its expansion amid persistent exchange rate fluctuations. Therefore, an in-depth investigation into its exchange rate risk management practices carries important practical implications.

This study collects data and documentation on Evergrande's foreign currency debt financing to analyze the status quo, deficiencies, and effectiveness of its exchange rate risk management. It also reviews domestic and international literature to draw on advanced theories and empirical findings while identifying research gaps.

2. Basic Concepts and Theoretical Foundations

2.1 Basic Concepts

2.1.1 Foreign Currency Debt Financing and Exchange Rate Risk

Exchange rate risk refers to the potential for economic losses arising from unanticipated exchange rate movements during cross-border transactions and foreign currency holdings [1]. It is typically categorized into three types: transaction risk, translation risk, and economic risk. For multinational and export-oriented firms, appropriate foreign currency borrowing can effectively hedge exchange rate risks associated with foreign assets. However, for firms with net foreign currency liabilities, foreign currency debt primarily increases exposure to transaction and translation risks [2].

2.1.2 Exchange Rate Risk Management

The selection of financial derivatives depends closely on the maturity structure of risk exposures. Forwards, options, and similar instruments are commonly used for short-term export risk hedging, whereas foreign currency debt and currency swaps are more suitable for long-term exchange rate risk control. Operational hedging achieves risk reduction through operational and strategic adjustments, while financial hedging relies on derivatives or structural arrangements of foreign currency debt [3].

Existing literature provides limited discussion of standardized practices for risk management in foreign currency debt financing. In practice, hedging remains the most widely adopted approach for managing exchange rate risks associated with foreign currency liabilities.

2.2 Theoretical Foundations

2.2.1 Purchasing Power Parity (PPP) Theory

Purchasing power parity theory states that exchange rates are determined by the relative purchasing power of two currencies, and long-run exchange rates tend to reflect inflation differentials [4]. Enterprises can use inflation differentials to forecast long-term exchange rate trends, match foreign currency debt maturities with hedging horizons, and reduce long-term exchange rate mismatch risks.

3. Current Status of Foreign Currency Debt Financing Risk Management of Evergrande Group

3.1 Introduction to Evergrande Group

Founded in Guangzhou in 1996 and headquartered in Guangzhou, Guangdong Province, Evergrande Group is a Fortune Global 500 conglomerate. It takes livelihood-oriented residential real estate as its core business, cultural tourism and health preservation as its coordinated sectors, and new energy vehicles as its leading industry [5]. Established by Xu Jiayin, the company initially focused on real estate development. With flexible market-oriented operation strategies and accurate judgment of market demand, Evergrande quickly secured a stable position in the Guangzhou real estate market and gradually realized diversified expansion of its business landscape.

3.2 Foreign Currency Debt Financing

3.2.1 Motivations and Advantages of Foreign Currency Debt Financing

The real estate industry is a typical capital-intensive sector, and overseas financing is a common approach for enterprises to broaden funding sources [6]. After the Central Economic Work Conference clearly put forward the principle that “houses are for living in, not for speculation” in 2016, relevant regulatory authorities successively issued policies to strengthen supervision over real estate enterprise financing and individual housing loans. Against this backdrop, domestic external financing channels for real estate enterprises tightened continuously, and internal funding channels mainly relying on sales collections were also restricted. As a result, enterprises in the industry stepped up overseas financing efforts .

With the continuous tightening of domestic regulation on the real estate industry, overseas bond issuance by real estate enterprises has also been regulated. In July 2019, the National Development and Reform Commission issued Document No. 778 (2019), stipulating that overseas debts issued by real estate enterprises could only be used to replace medium and long-term overseas debts maturing within the next year. This indicates that the scale of foreign currency debts in the real estate industry will hardly expand in the short term.

After the global financial crisis, economies such as the United States and the United Kingdom cut benchmark interest rates to near zero to boost economic recovery and implemented quantitative easing policies. These measures drove down medium and long-term treasury bond rates, with yields significantly lower than those in emerging economies including China, leading to large capital inflows into emerging markets. Since 2021, the China-US interest rate spread has remained high in most periods. The low US dollar interest rate environment has created favorable conditions for Evergrande Group to leverage overseas financing channels and borrow low-cost funds.

Promoting international layout is an important part of Evergrande Group's long-term development strategy. In 2017, Evergrande officially entered the overseas market by cooperating with high-quality overseas real estate enterprises. Later, it continuously expanded overseas businesses by acquiring overseas companies and projects, carrying out real estate development and asset management in many cities around the world. To support the implementation of the international strategy and the advancement of cross-border mergers and acquisitions of domestic real estate businesses, the Group's demand for foreign currency funds rose sharply. To this end, Evergrande actively integrated high-quality overseas resources, fully utilized overseas financing platforms to issue bonds, and continuously expanded the scale of overseas financing.

3.2.2 Scale and Structure of Foreign Currency Debt Financing

Evergrande's foreign currency debt financing mainly relies on bank loans and overseas bond issuance, with long-term loans as the main component of foreign currency borrowings. The US dollar and Hong Kong dollar are the core borrowing currencies. Data from 2023 show that US dollar borrowings and Hong Kong dollar borrowings accounted for 49.9% and 47.11% of the Group's total foreign currency borrowings, respectively. In addition, due to overseas real estate projects in Singapore and the UK, Evergrande also has a small amount of liabilities in Singapore dollars and British pounds. All overseas bonds issued by the Group are medium and long-term, with US dollar bonds as the dominant type and Hong Kong dollar bonds as the secondary type.

3.3 Exchange Rate Risk Management Model

3.3.1 Exchange Rate Risk Management Strategies and Systems

Exchange rate risk management is a highly professional task that requires professionals with solid financial knowledge, rich market experience and keen risk awareness. Evergrande Group has improved the exchange rate risk management capabilities of existing financial staff through internal training, external exchanges and university cooperation. At the same time, it has actively introduced high-end talents with international financial market experience to strengthen the exchange rate risk management team, providing intellectual support for formulating scientific and reasonable exchange rate risk management strategies.

To ensure the effective implementation of exchange rate risk management strategies, Evergrande Group has established a performance appraisal mechanism compatible with exchange rate risk management, incorporating the effectiveness of exchange rate risk management into the performance appraisal indicator system of relevant departments and personnel . Through reasonable performance incentives, employees are guided to actively participate in exchange rate risk management, raise their awareness and attention to exchange rate risks, and encouraged to take initiatives to reduce exchange rate risks in daily work, forming a sound risk management culture.

3.3.2 Application of Exchange Rate Risk Management Tools

To achieve optimal exchange rate risk management results, Evergrande Group does not rely on a single tool. Instead, it comprehensively uses a variety of tools to build a hedging portfolio strategy based on its own risk profile, market expectations and cost-benefit factors. While signing forward foreign exchange contracts to lock in the basic exchange rate, the Group purchases foreign exchange options to obtain opportunities for additional benefits from favorable market changes. This combination strategy increases profit flexibility on the premise of ensuring basic control over exchange rate risks. When market exchange rates fluctuate sharply, Evergrande increases the proportion of financial derivatives, raises the purchase volume of foreign exchange options or adjusts the scale of currency swaps. When market exchange rates tend to be stable, it appropriately reduces the use of financial derivatives to lower risk management costs.

4. Analysis of Exchange Rate Risks in Foreign Currency Debt Financing of Evergrande Group

4.1 Types of Exchange Rate Risks

Evergrande Group mainly faces two types of exchange rate risks: transaction risk and translation risk. Transaction risk refers to the possibility of economic losses arising from exchange rate differences between the transaction date and the actual settlement date after an enterprise forms foreign currency claims or undertakes foreign currency debts. In 2024, foreign currency borrowings accounted for 20.47% of Evergrande's total interest-bearing liabilities. The debt currencies were highly concentrated in US dollars and Hong Kong dollars, making risk hedging difficult due to the single currency structure.Coupled with the complex global economic environment and increasing uncertainty in RMB exchange rate fluctuations, Evergrande faced significant transaction risk pressure . Between 2016 and 2024, Evergrande incurred large exchange losses in four years due to sharp fluctuations in US dollar and Hong Kong dollar exchange rates. Although the impact of exchange gains and losses on net profit was generally controllable, the transaction risk caused by foreign currency liabilities cannot be ignored.

Some overseas subsidiaries of Evergrande use non-RMB currencies as their functional currencies. When preparing consolidated financial statements, foreign currency-denominated items must be translated into RMB, resulting in translation risk. This risk is reflected in the foreign currency translation difference under other comprehensive income in the statements. Since the foreign currency translation difference is only an accounting adjustment item, not a real operating profit or loss, and has no substantial impact on corporate cash flow and net profit, Evergrande's risk hedging operations mainly focus on controlling transaction risk.

4.2 Exchange Rate Risk Exposure

According to the 2022 earnings announcement, as of December 31, 2022, China Evergrande's interest-bearing liabilities reached 612.391 billion yuan, accounting for 25.12% of total liabilities. Short-term interest-bearing liabilities accounted for 24.08% of total debt, a decrease of 14.53% from the end of 2020. However, the company's available cash was only 14.305 billion yuan (including restricted funds). This indicates that Evergrande Real Estate fell into a debt repayment dilemma between 2021 and 2022, shifting its business focus from expansion to "ensuring housing delivery". The disclosed long-term borrowings were mainly unmatured bonds and US dollar senior notes. In the interest-bearing liability structure, RMB-denominated and US dollar-denominated debts accounted for 75.95% and 21.23% respectively, forming obvious exchange rate risk exposure.

4.3 Impact Analysis of Exchange Rate Fluctuations

The depreciation of RMB increased the RMB-equivalent amount of Evergrande's foreign currency debts, exacerbating its debt repayment pressure. For example, Evergrande suffered exchange losses of 9.59 billion yuan in 2022, which directly increased financial expenses and weakened debt solvency . Due to the high proportion of foreign currency debts, Evergrande's short-term debt repayment pressure was particularly prominent. In 2022, short-term interest-bearing liabilities accounted for 24.08% of total debt, while cash was only 14.305 billion yuan, far insufficient to cover short-term debts. Translation risk caused by exchange rate fluctuations increased the amount of foreign currency liabilities on Evergrande's balance sheet, further deteriorating its financial position. This not only affected Evergrande's financial indicators but also triggered market concerns about its financial health.

5. Problems and Suggestions on Exchange Rate Risk Management of Foreign Currency Debt Financing of Evergrande Group

5.1 Existing Problems in Exchange Rate Risk Management

5.1.1 Exchange Rate Risk Management System

Evergrande Group lacks a unified internal awareness of exchange rate risk management, leading to inconsistencies in the implementation of risk management measures among different departments and impairing overall effectiveness. The exchange rate risk management system has not been updated in a timely manner to adapt to changing market conditions and business development, making it unable to effectively address new risk challenges. Exchange rate risk management involves finance,

business, legal affairs and other departments, but the coordination mechanism among departments is imperfect, resulting in poor information transmission and failure to form an effective risk management synergy.

5.1.2 Application of Exchange Rate Risk Management Tools

Evergrande Group lacks the initiative to use hedging tools in exchange rate risk management and failed to hedge exchange rate risks, leading to huge exchange losses when RMB depreciated. When selecting hedging instruments, Evergrande may rely too heavily on a single tool such as forward foreign exchange contracts, without fully considering the combined use of other instruments including foreign exchange options and foreign exchange swaps, reducing risk management effectiveness. Non-standard operations may exist in hedging practices, such as failure to comply with hedging accounting standards for bookkeeping, leading to inaccurate reflection of hedging effects in financial statements. The adjustment of asset-liability matching strategies takes a long time and incurs high costs, which may affect corporate capital liquidity. Evergrande faces great pressure in capital liquidity management and struggles to adjust its asset-liability structure timely to cope with exchange rate fluctuations.

5.2 Improvement Suggestions

5.2.1 Improve the Exchange Rate Risk Management System

The core objective of Evergrande Group's foreign exchange risk management is to minimize the potential negative impacts of exchange rate fluctuations on the company's financial position. To achieve this goal, the Group has established a sound exchange rate risk management system with the following principles: first, take prevention and control of exchange rate risks as the fundamental orientation and strictly prohibit speculative operations; second, optimize the matching of investment and financing scale and currency structure, and rationally apply hedging instruments; third, select appropriate hedging tools based on real business background, combined with transaction currency, amount and maturity; fourth, continuously track the dynamics of the foreign exchange market and flexibly adjust hedging strategies.

5.2.2 Optimize the Application of Exchange Rate Risk Management Tools

To mitigate the negative impacts of exchange rate fluctuations and reduce the repayment risk of foreign currency debts, Evergrande Group and its overseas subsidiaries have long taken the initiative to carry out foreign exchange hedging operations. When selecting financial contracts, the Group comprehensively considers factors such as exchange rate trends, transaction currencies, contract maturities and transaction scales to choose appropriate financial instruments. The foreign exchange derivatives actually used mainly include non-deliverable forwards (NDF), deliverable forwards (DF) and cross-currency swaps (CCS). The use of these derivatives has effectively reduced Evergrande's foreign exchange risk exposure and alleviated the adverse impacts of exchange rate fluctuations on the company's net profit.

6. Conclusion

6.1 Research Conclusions

Taking Evergrande Group as a case, this study conducts an in-depth analysis of the current status and existing problems of exchange rate risk management in foreign currency debt financing, and puts forward corresponding optimization strategies. The study finds that Evergrande has various deficiencies in exchange rate risk management, which not only reduce the effectiveness of risk management but also increase corporate financial and operational risks. Specifically, shortcomings exist in the implementation of risk management systems, the application of risk management tools and the evaluation of management effectiveness. These problems make it difficult for Evergrande to effectively respond to exchange rate fluctuations, resulting in significant impacts on its financial position and operating performance. Targeted optimization strategies are proposed to help Evergrande improve its exchange rate risk management system, optimize the use of risk management tools and refine the management effectiveness evaluation system, thereby enhancing its overall risk management capacity. Furthermore, this study provides useful references for other enterprises engaged in foreign currency debt financing, helping them operate steadily in the complex international financial market and improving the overall risk management level of Chinese enterprises in international financing.

6.2 Deficiencies and Prospects

This study mainly relies on the financial reports and relevant data publicly disclosed by Evergrande Group, which may have certain limitations and cannot fully reflect all details of the Group's exchange rate risk management practices. In addition, due to the timeliness and integrity of data, some analysis results may lack precision. This study mainly adopts case analysis and literature research methods. Although these methods can reveal the current status and problems of Evergrande's exchange rate risk management to a certain extent, the connection between different methods may be insufficient in integrated application, leading to weak persuasiveness of some analysis results.

Future research can further expand data sources. In addition to publicly disclosed financial reports, internal data and information can be obtained through field investigations and interviews to comprehensively understand Evergrande's exchange rate risk management practices. Big data analysis technology can be used to deeply mine and analyze Evergrande's risk management data, improving the accuracy and reliability of research results. Future studies can conduct multi-case comparative research by selecting enterprises in different industries and of different scales, summarizing experiences and lessons in exchange rate risk management of foreign currency debt financing for different types of enterprises, and providing more targeted references for other enterprises.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Batbold, B., Kikuchi, K., & Kusuda, K. (2025). Strategic international asset allocation under a quadratic model with exchange rate and inflation-deflation risks. Decisions in Economics and Finance, 1–34.

[2] Cao, Y., & Wang, Q. (2025). Export factoring in global supply chain: Information updating and convertible contracts. Journal of the Operational Research Society, 76, 1104–1124.

[3] Castro, L. D., Claudio, F., & Arthur, R. (2025). How to deal with exchange rate risk in infrastructure and other long-lived projects. Journal of Public Economic Theory, 27, e70003.

[4] Egozcue, M., & García, L. F. (2024). Optimal bounds and practical insights: Cantelli's inequality revisited. Journal of Statistical Theory and Practice, 18, 59.

[5] Huong, N. T. T. (2025). Impact of exchange rate risks on trade flows: Does regional integration in production matter? Journal of International Commerce, Economics and Policy.

[6] Jacoby, G., Liao, R. C., Wang, Y., & Wu, Z. (2025). An intertemporal international asset pricing model: Theory and evidence. Journal of International Financial Markets, Institutions & Money, 102, 102162.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

A Comparison of Global Aging Trends: Exploring the Path from "Healthy Aging" to "Active Aging" - Case Studies Based on Japan, Northern Europe and China

Jiarui Hu*, Mengyuan Fang, Jiahui Jiang, Haojun Huan, Zhengxiu Feng

School of Public Administration, Shandong Technology and Business University, Yantai, Shandong, 264005, China

Corresponding author: Jiarui Hu, 3194343810@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: The global population structure is undergoing a profound aging transformation, driving the concepts of "healthy aging" and "active aging" to gradually evolve from focusing on the maintenance of individual physical health to a comprehensive framework emphasizing social participation, equal opportunities, and the development of the entire life cycle. This article aims to deeply analyze how cultural background, institutional arrangements and social structure shape the differentiated paths for each country to deal with aging by comparing three typical cases: Japan (ultra-aged society and cultural tradition integration), Northern Europe (high welfare system and universalist design), and China (rapid aging and aging before getting rich). The core research finds that the governance model for aging is deeply influenced by cultural values, the maturity of the welfare system, and the intergenerational relationship structure. Among them, the construction of social support networks, the collaboration of multiple subjects, and the integration of life-cycle policies are key variables. The implications for China lie in the fact that it is necessary to base itself on the local filial piety culture and the stage of social and economic development. While strengthening the family-based elderly care function, it is also necessary to accelerate the construction of a multi-level elderly care security system, promote the construction of an age-friendly environment, and actively explore aging response plans that are culturally adaptable, institutional resilient and social innovative.

Keywords: Population Aging; Healthy Aging; Active Aging; Comparative Study; Japan; Northern Europe; China

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1407

1. Main Research Conclusions

With the acceleration of global population aging, national response strategies are undergoing a subtle transformation. The previously health-focused "healthy aging" model is gradually being replaced by "active aging," which emphasizes encouraging social participation, promoting employment, and enhancing capabilities[1][2]. This study examines Japan, the Nordic region, and China as case studies to conduct an in-depth analysis of their similarities and differences in transition pathways, policy tool applications, and the development of social support systems, leading to relevant conclusions.

1.1 The Unique Path and Determining Factors of the Transition from "Healthy" to "Active" Aging in the Three Kingdoms Period

In addressing aging challenges, Japan has developed an effective model through comprehensive legislation. The country enacted key laws including the Elderly Employment Stabilization Act and the Lifelong Active Society Promotion Act,

mandating corporate retirement age extensions and providing reemployment support[3][4][5]. This established a robust framework characterized by "government-led legislation, corporate compliance, and multi-stakeholder collaboration." Japan demonstrates strategic foresight in employment guidance by prioritizing senior citizens in experience-intensive sectors like services and consulting, effectively harnessing "the value of accumulated expertise."[5] The nation maintains a comprehensive Silver Human Resources Center network offering integrated services such as job placement, vocational training, and career counseling to facilitate reintegration into the workforce. Furthermore, Japan's deeply rooted work ethic fosters a culture where many seniors view employment as vital to life fulfillment. Retirement without work often triggers feelings of social exclusion and unrealized potential, compelling them to continue working—not merely for financial gain but also to maintain social connections and achieve personal fulfillment[4].

For Nordic countries, building upon the well-established foundation of a "high taxation-high welfare" system, seniors have been redefined as "lifelong learners" and "active participants."[6][7][8] The implementation strategy relies on three key pillars: a universal lifelong education system, flexible retirement policies, and community-integrated services, with particular emphasis on the concept of "work-life balance." At the governmental level, tax incentives encourage businesses to actively employ older adults, while community centers provide skills training, volunteer opportunities, and social support services. This has formed a flexible integration model characterized by "state guarantees as solid support, community engagement as practical foundation, and individual autonomy as intrinsic motivation."[6][8] Notably, the Nordic model places great importance on elderly mental health and social connections, actively promoting social integration through intergenerational cohabitation programs and community cultural activities—practices deeply rooted in the Nordic philosophy of equality for all citizens[7].

Currently, China's population development has entered a completely new phase, characterized by declining birth rates, an aging population, and significant regional disparities in population growth and decline. Within traditional Chinese family values, "caring for the elderly" holds a prominent position, while the concept of "utilizing the elderly" is rarely emphasized. This deeply ingrained mindset results in generally low societal and individual acceptance of elderly reemployment[1][9][10]. For most seniors, although they recognize the need for "care," they fundamentally wish to retain their personal value. Moreover, societal pressures act like an invisible shackle, preventing them and their children from typically encouraging elderly individuals to seek employment elsewhere. Additionally, the gap between urban and rural areas is pronounced: rural seniors predominantly engage in agricultural labor, while urban seniors face a mismatch between their reemployment aspirations and available opportunities. In the information age, Chinese seniors have increasingly limited access to information, hindering their full integration into this rapidly evolving era. Their adoption of new technologies is relatively slow, and their children are often preoccupied with livelihoods, leaving little time for support. China's silver economy remains in its early stages of development, confronting dual challenges of "aging before prosperity" and "aging without adequate preparation."[11][12] The labor market has yet to establish flexible mechanisms to accommodate the elderly, and social security systems—such as pensions and healthcare—are primarily focused on meeting basic needs. Vocational training resources are predominantly allocated to younger generations, leaving seniors with limited access to adequate training opportunities[13]. Compared with countries like Japan and Northern Europe, China still lacks a well-established system for vocational training for the elderly, with severely inadequate incentives for enterprises to employ older workers and a lack of systematic legal safeguards. There are few easy and engaging employment opportunities suitable for seniors. Even when some older adults break free from social pressures, they are plagued by questions such as "what jobs are available" and "how to secure employment opportunities." Clearly, the path to re-employment for the elderly is fraught with obstacles. Despite many seniors possessing work capabilities and a desire to re-enter the workforce, they often have to abandon their aspirations. We need to focus on both mindset shifts and institutional measures, integrating the concept of active aging and healthy aging into the entire process of economic and social development[1][2]. However, directly transplanting the Japanese or Nordic models may lead to "institutional rejection," necessitating dual efforts in policy localization and social mindset guidance[14].

2. Research Gaps and Future Prospects

2.1 Shortage of Research, Inadequate Research, Lack of Research

This study faces challenges in data depth and microscopic mechanism analysis. Current research predominantly focuses on

macro policies and statistical descriptions, lacking in-depth exploration of micro-level mechanisms such as the formation of elderly re-employment decisions, underlying motivations for enterprises to hire seniors, and the operational dynamics of community support networks. Particularly regarding the psychological transition of older adults from "health-oriented aging" to "active participation," as well as the specific mechanisms through which family support systems function, more in-depth qualitative research is urgently needed[9][13].

The study lacks sufficient exploration of internal disparities within Northern Europe. By treating the region as a homogeneous entity, it fails to meticulously differentiate subtle variations in elderly employment policies and community service models among countries like Sweden, Denmark, and Norway. In reality, Nordic nations exhibit distinct differences in pension systems, employment subsidy standards for seniors, and approaches to community service delivery. These variations significantly impact policy effectiveness, warranting in-depth investigation[6][8].

The study lacks in-depth analysis of regional disparities in China regarding the differentiation trends in aging response strategies among eastern, central, and western regions as well as urban and rural areas. Given China's vast territory and significant differences in economic development, aging levels, and cultural traditions across regions, conducting more detailed regional comparative studies is essential[9][10].

The study exhibits limitations at the data level, primarily relying on cross-sectional data or short-term policy evaluations, lacking long-term tracking of the transition process from 'healthy aging' to 'active aging' among the elderly population. This results in difficulties in accurately assessing the long-term efficacy and sustainability of relevant policies[1][2].

2.2 Future Research Direction

Further in-depth comparative studies should be conducted between urban and rural areas as well as different regions, focusing on China's various regions such as the Yangtze River Delta, Pearl River Delta, and rural areas in central and western China, to explore their differentiated approaches in elderly human resource development and community support system construction. Particular attention should be paid to how rural elderly can achieve "productive aging" amidst the wave of agricultural modernization, while also helping younger urban elderly break through the "glass ceiling of re-employment" and realize their own value[9][12].

It is necessary to conduct cross-border impact assessments of key policies. Specifically, the policy effectiveness of Japan's "Silver Human Resources Center," Northern Europe's "flexible retirement system," and China's "long-term care insurance pilot program" in enhancing elderly social participation can be compared. Quasi-experimental designs or case comparison methods can be used to accurately evaluate the cost-effectiveness and social impacts of different policy tools[4][5][14].

To conduct in-depth research in the fields of cultural psychology and social identity, cross-cultural surveys should be implemented to analyze differences between China and Japan, as well as China and Europe, particularly in aspects such as "perception of elderly values" and "stigma surrounding re-employment." These findings will provide a solid foundation for policy dissemination and social advocacy. Social psychology theories suggest leveraging narrative transformation and role model promotion to reshape societal perceptions of elderly participation in social activities—a critical priority that demands immediate attention[1][14].

The research should focus on multi-stakeholder collaborative mechanisms, conducting an in-depth analysis of the roles and synergistic approaches of governments, enterprises, communities, and families in advancing the "active aging" initiative[15]. Particular attention should be given to exploring how enterprises balance economic benefits with social responsibilities, and how communities integrate various resources to create a supportive environment for the elderly[8][11][12].

Against the backdrop of an aging society, in-depth research on intergenerational relationships and family support has become imperative. It is essential to examine the evolving trends in intergenerational dynamics and the impact of family support systems on elderly social participation. Particular attention should be paid to studying children's attitudes toward parental re-employment and their influencing factors, thereby providing solid evidence for family policy formulation[1][9].

It is imperative to establish an assessment indicator system for "active aging" that aligns with China's national conditions. This system must encompass multiple dimensions such as health, participation, and protection to provide scientific and precise tools for policy formulation and efficacy evaluation[2][12].

Funding

This article is from 2025 College Students' Innovation and Entrepreneurship Training Program Funded Project, Project No. X2025116880098.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Yang, J. H. (2023). From “healthy aging” to “active aging”: Reflections and prospects on the China path. Population Research, 47(4), 45–58.

[2] World Health Organization. (2025). Decade of healthy ageing: Baseline report 2025. Geneva: WHO.

[3] Ministry of Health, Labour and Welfare of Japan. (2025). 2024 edition of the white paper on an aging society: Employment and social participation in the 100-year life era. Tokyo: Nikkei Printing.

[4] Suzuki, M., & Tanaka, K. (2025). Institutional embeddedness and senior employment: A neo-institutional analysis of Japan’s “Society 5.0 for All Ages”. Journal of Aging Studies, 68, 101–118.

[5] Hiyoshi, A., & Kondo, N. (2022). The Japanese model of elderly employment: Legislation, corporate response, and social outcomes. Journal of Aging & Social Policy, 34(3), 312–330.

[6] Rostgaard, T., & Jacobsen, F. (2025). Beyond the welfare state: The role of civil society in promoting late-life participation in Scandinavia. Voluntas, 36(2), 245–261.

[7] Swedish Ministry of Health and Social Affairs. (2025). The Nordic model for active ageing 2025: Mid-term evaluation report. Stockholm: Government Offices.

[8] Rostgaard, T., & Pavolini, E. (2023). Nordic ageing policies: Between universalism and activation. Social Policy & Administration, 57(2), 221–237.

[9] Du, P., & Yang, G. (2025). Regional practice atlas of active aging in China (2020–2025): Comparative analysis based on multiple cases. China Social Sciences, (4), 112–133+207.

[10] National Health Commission & Office of the National Working Committee on Aging. (2024). 2023 national bulletin on aging development. Beijing: China Statistics Press.

[11] National Development and Reform Commission & Ministry of Human Resources and Social Security. (2025). First-year assessment report on the implementation of the silver economy promotion regulations (2025). Beijing: China Development Press.

[12] Tsinghua University Research Center for Aging Society. (2024). China active aging development report (2023). Beijing: Social Sciences Academic Press.

[13] Li, J., & Xue, K. W. (2025). Digital inclusive design and high-quality employment for younger seniors: An empirical study based on human-computer interaction experiments. Management World, 41(7), 88–104.

[14] Chen, S., & Walker, A. (2025). Applying the COM-B model to active ageing policies in East Asia: A comparative study of China, Japan, and South Korea. The Gerontologist, 65(3), 456–470.

[15] Esping-Andersen, G. (2025). The three worlds of welfare capitalism (30th anniversary ed.). Princeton: Princeton University Press.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

The Impact of Market-Incentivized Environmental Regulation on Green Innovation in Energy Enterprises

Yi Ge*

School of Management, Xi'an Polytechnic University, Shaanxi, Xi'an, 710600, China

Corresponding author: Yi Ge, 18852394888@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Against the backdrop of the "dual carbon" goals and the green transformation of the energy industry, whether market-incentivized environmental regulation can effectively promote green innovation in energy enterprises has become a key issue. This paper takes listed energy companies on the Shanghai and Shenzhen A-share markets from 2012 to 2022 as the sample, and based on data such as pollutant discharge fee amounts, empirically examines the impact of market-incentivized environmental regulation on green innovation in energy enterprises and its regional heterogeneity. The findings reveal that market-incentivized environmental regulation significantly promotes green innovation in energy enterprises overall, verifying the applicability of the Porter Hypothesis in the energy industry. This promotional effect is significant in non-coastal regions but not significant in coastal regions, reflecting differences in regional economic development stages, market maturity, and policy coordination capabilities. Accordingly, this paper proposes policy recommendations including regionally differentiated policy design, strengthening foundational support in non-coastal areas, and deepening market mechanism coordination in coastal areas, providing empirical evidence and policy references for optimizing environmental regulation policies and promoting the green transformation of energy enterprises.

Keywords: Market-based Environmental Regulation; Green Innovation of Energy Enterprises

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1415

Introduction

In the context of deepening global climate governance and the continuous advancement of ecological civilization construction, energy enterprises, as foundational industries characterized by high pollution and high energy consumption, are facing an urgent need to transition towards green, low-carbon, intensive, and high-quality development. Achieving the "dual carbon" goals requires the construction of a clean, low-carbon modern energy system, which has spurred increasingly stringent and diversified environmental regulation policies[1]. Among these, market-incentivized environmental regulations, represented by emissions trading, green technology subsidies, and environmental taxes, are becoming core components of the policy toolkit. The essence of such tools lies in internalizing external environmental costs through price transmission mechanisms, providing economic incentives or penalties for corporate production activities to achieve pollution reduction. Their core characteristic is that they attempt to convey clear green development guidance signals and economic incentives to high-emission, high-energy-consumption energy enterprises by altering market factor prices and competition rules, thereby influencing their resource allocation and innovation decisions.

However, amidst this deep transformation process in the energy industry, a core and insufficiently answered key question is: Can current market-incentivized environmental regulation effectively promote green innovation in energy enterprises? More importantly, are there significant differences in this promotional effect under different contexts? Observations from reality add complexity to this inquiry: as the intensity of such regulations increases, green innovation activities in some energy enterprises are effectively stimulated, while others respond slowly or even experience innovation stagnation. This divergence forces us to confront a core question: Is the net effect of market-incentivized environmental regulation on green innovation in energy enterprises a powerful driving force, or does it vary depending on conditions?

Clarifying this fundamental issue is not only crucial for the applicability of environmental regulation theory in the context of energy transition but also directly determines whether policy design can be precisely targeted to effectively guide this key industry in breaking through green innovation bottlenecks, ultimately serving the realization of the national "dual carbon" strategic goals.

1. Mechanism Analysis and Hypothesis Development

1.1 Market-Incentivized Environmental Regulation and Corporate Green Innovation

The core of market-incentivized environmental regulation lies in its incentive compatibility attribute. It internalizes environmental costs into corporate decision-making variables through economic levers, guiding enterprises to optimize their environmental behavior. More importantly, market-incentivized environmental regulation itself is a powerful policy signal, conveying to enterprises the government's determination for green transformation and the core value orientation of environmental performance, reshaping enterprises' long-term expectations and strategic priorities for green innovation, and stimulating endogenous motivation. According to the Porter Hypothesis, appropriately designed environmental regulations can stimulate corporate green technology innovation through innovation compensation effects, enhancing long-term competitiveness. Therefore, market-incentivized environmental regulation is expected to become an important driving force for green innovation in energy enterprises. Accordingly, the following hypothesis is proposed:

H1: Market-incentivized environmental regulation has a significant promotional effect on green innovation in energy enterprises.

2. Research Design

2.1 Sample Selection and Data Sources

This study takes listed companies in the energy industry on the Shanghai and Shenzhen A-share markets from 2012 to 2022 as the initial sample, constructing an empirical analysis framework incorporating multi-source data. The specific data collection and processing procedures are as follows: Green patent application data originates from the CNRDS patent database. The market-incentivized environmental regulation indicator is constructed by quantifying relevant policy provisions from the China Environmental Statistics Yearbook, China Industry Statistics Yearbook, China Environment Yearbook, and China Tax Yearbook. To ensure data quality, the following screening procedures were applied sequentially: (1) Excluding companies with ST and *ST special treatment status to eliminate interference from abnormal operations; (2) Removing observations with missing data or extreme outliers in control variables. After the above processing, 1,141 valid observations were obtained.

2.2 Variable Definitions

This study uses Green Innovation (GI) as the dependent variable. Following mainstream research paradigms[2,3], it is measured using the natural logarithm of one plus the number of green patent applications filed by the enterprise. This indicator is based on the CNRDS patent database, which comprehensively covers the application and authorization information of green patents for Chinese listed companies. Green patents require meeting high technological R&D thresholds and can objectively reflect the comprehensive innovation achievements of enterprises in product performance improvement and green technology breakthroughs, ensuring the stability and reliability of the data source.

The core explanatory variable focuses on Market-Incentivized Environmental Regulation (ERI). Its quantification method follows the design of Wu Lei et al.[4], using the amount of pollutant discharge fee levied as a proxy variable.

Drawing on existing research, the control variables selected in this paper are: Board Size (BDS), Proportion of Independent

Directors (IDP), Supervisory Board Size (SBS), Book-to-Market Ratio (BM), Price-to-Book Ratio (PB), and Firm Age (Age). The variable definitions are shown in Table 1

Table 1: Variable Definitions

Variable Name	Symbol	Variable Definition
Green Innovation	GL	Natural logarithm of one plus the number of enterprise green patent applications
Market-Incentivized Environmental Regulation	ERI	Natural logarithm of the pollutant discharge fee amount
Board Size	BDS	Natural logarithm of one plus the total number of board members
Proportion of Independent Directors	IDP	Proportion of external directors not involved in company operations on the board
Supervisory Board Size	SBS	Natural logarithm of one plus the number of supervisory board members overseeing management
Book-to-Market Ratio	BM	Ratio of company book value to market value
Price-to-Book Ratio	PB	Ratio of stock price per share to net assets per share
Firm Age	Age	Time span since the company's registration and establishment

3. Empirical Analysis

3.1 Descriptive Statistics

The descriptive statistics for the main variables show that the mean of GL is 0.709, which is significantly greater than the median of 0, with a minimum of 0 and a maximum of 6.922, indicating that the green innovation capability of most energy enterprises is below the average level and that there is a large internal disparity in green innovation capability among energy enterprises. The mean of market-incentivized environmental regulation is 2.06, with a median of 2.017, indicating that nearly half of the energy enterprises face relatively large pollutant discharge fees or environmental tax amounts. The maximum value of institutional shareholding ratio is 0.987 and the minimum is 0, indicating a large disparity in institutional shareholding ratios among energy enterprises.

Table 2: Descriptive Statistics

VarName	Obs	Mean	SD	Median	Min	Max
GL	1141	0.709	1.077	0.000	0.000	6.922
ERI	1141	2.060	0.676	2.017	0.251	3.608
BDS	1141	2.303	0.202	2.303	1.386	2.944
IDP	1141	37.008	4.779	35.290	25.000	66.670
SBS	1141	1.561	0.257	1.386	0.693	2.485
BM	1141	0.707	0.246	0.713	0.064	1.288
PB	1141	3.651	15.024	2.208	0.476	324.161
Age	1141	2.873	0.319	2.890	1.386	3.611

3.2 Regression Analysis

3.2.1 Model Specification

To analyze the impact of market-incentivized environmental regulation on green innovation in energy enterprises, this paper constructs the following two-way fixed effects model:

$$GL=\beta_0+\varphi ERI_{it}+\gamma X_{it}+\alpha_i+\delta_t+\varepsilon_{it} \quad (1)$$

Where ERI represents market-incentivized environmental regulation, GL represents green innovation in energy enterprises, X represents the control variables;andrepresent individual and time fixed effects, is the random error term, i denotes province, and t denotes year.

3.2.2 Direct Effect Analysis

To verify the direct effect of market-incentivized environmental regulation on green innovation in energy enterprises, a direct effect analysis was conducted. The results show that in Model 1, which does not include control variables, the regression coefficient for market-incentivized environmental regulation (ERI) is positive at the 5% significance level, indicating that for each unit increase in market-incentivized environmental regulation, green innovation in energy enterprises increases by 0.206 units. Model 2 presents the regression results after adding control variables, where the impact of market-incentivized environmental regulation on green innovation in energy enterprises remains positive. Hypothesis H1, that market-incentivized environmental regulation has a significant promotional effect on green innovation in energy enterprises, is supported.

Table 3: Baseline Regression Results

	(1)	(2)
	GL	**GL**
ERI	0.206**	0.215**
	[2.37]	[2.47]
BDS		1.174***
		[4.25]
IDP		0.0197**
		[2.58]
SBS		0.157
		[0.58]
BM		-0.0112
		[-0.08]
PB		-0.00180
		[-0.40]
Age		0.637
		[1.41]
Individual Fixed Effects	0.286	-5.231***
Time Fixed Effects	[1.59]	[-3.34]
N	1114	1114
R^2	0.710	0.717

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

3.2.3 Robustness Tests

(1) Using Lagged Explanatory Variables. Market-incentivized environmental regulation may have a lagged effect on green innovation in energy enterprises. Therefore, this paper lags the explanatory variable by one and two periods and performs regression analysis separately, as detailed in columns (1) and (2) of Table 4. The coefficients for ERI lagged by one and two periods are both positive at the 5% significance level. The results indicate that market-incentivized environmental regulation has a significant promotional effect on green innovation in energy enterprises, consistent with the baseline regression results, confirming the robustness of the findings.

(2) Sample Data Screening. Extreme values may affect the baseline regression results. Therefore, the sample was winsorized at 1% and 5% levels, and the regressions were re-run, as shown in columns (3) and (4) of Table 4. After excluding extreme values, the regression coefficient for ERI remains significantly positive at the 5% level, consistent with the baseline regression results.

Table 4: Robustness Test

	(1)	(2)	(3)	(4)
	GL	**GL**	**GL**	**GL**
L.ERI	0.258**			
	[2.54]			
L2.ERI		0.310**		
		[2.38]		
L.GL				
ERI			0.205**	0.194**
			[2.32]	[2.23]
BDS	1.287***	1.401***	1.158***	1.253***
	[3.86]	[3.86]	[4.17]	[4.85]
IDP	0.0219**	0.0222**	0.0177**	0.0216**
	[2.40]	[2.19]	[2.10]	[2.39]
SBS	0.125	0.190	-0.0791	0.0227
	[0.41]	[0.58]	[-0.25]	[0.07]
BM	-0.0850	0.0201	-0.116	-0.164
	[-0.46]	[0.09]	[-0.64]	[-0.87]
PB	-0.00286	-0.00376	-0.0207	-0.0269
	[-0.28]	[-0.59]	[-1.03]	[-1.05]
Age	0.562	0.993	1.237**	0.883**
	[0.92]	[1.20]	[2.36]	[2.08]
Individual Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
N	909	758	1114	1114
R^2	0.720	0.682	0.690	0.642

3.2.4 Heterogeneity Test

Coastal vs. Non-Coastal Regions

Due to differences in environmental policies and economic development across regions, the influence mechanism on green innovation in energy enterprises also exhibits certain regional characteristics. Based on this, this paper further divides the sample into coastal and non-coastal regions to explore the differences in regional effects. Coastal regions include Tianjin, Hebei, Liaoning, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, and Guangdong; the remaining areas are classified as non-coastal regions. See Table 5, where column (1) shows coastal regions and column (2) shows non-coastal regions.

Table 5: Heterogeneity Test

	(1)	(2)
	GL	**GL**
ERI	0.0782	0.371***
	[0.69]	[2.72]
BDS	1.614***	0.380
	[4.75]	[0.86]
IDP	0.0306**	0.00726
	[2.49]	[0.67]
SBS	-0.0178	0.423
	[-0.05]	[1.14]
BM	-0.257	0.0643
	[-0.89]	[0.34]
PB	-0.0223	0.000188
	[-0.85]	[0.04]
Age	-1.874	0.857*
	[-1.61]	[1.70]
N	404	710
R^2	0.643	0.748

The insignificant impact of market-incentivized environmental regulation on green innovation in energy enterprises in coastal regions may stem from the particularities of their economic structure and policy environment. On one hand, coastal regions underwent industrial upgrades earlier, and energy enterprises already possess certain green technology reserves, leaving limited marginal incentive space for policy. On the other hand, market mechanisms in coastal regions are mature, and enterprises face diverse external pressures, diluting the policy signals of environmental regulation. Additionally, coastal enterprises have ample R&D resources and a tendency towards risk aversion, making them less sensitive to short-term policy fluctuations and more reliant on long-term technological accumulation rather than immediate policy responses, further weakening statistical significance.

The significant effect in non-coastal regions is closely related to regional latecomer advantages and policy gaps, meaning market-incentivized environmental regulation yields policy dividends for them. Non-coastal energy enterprises typically have lower technological starting points and stronger path dependence. Market-incentivized policies directly break the “high pollution - low cost” equilibrium, creating significant expectations for innovation returns. Simultaneously, local governments in non-coastal areas often strengthen policy coordination through targeted support, lowering the barriers to enterprise transformation. Coupled with less competitive pressure in the region, policy signals become a core trigger for innovation, and enterprises can more easily achieve green catching-up through rapid application of mature technologies, making the policy effects significantly evident both statistically and practically.

4. Research Conclusions

This paper takes listed energy companies on the Shanghai and Shenzhen A-share markets from 2012 to 2022 as the research sample, empirically examines the direct impact of market-incentivized environmental regulation on green innovation in energy enterprises, and focuses on the heterogeneous effects between coastal and non-coastal regions. The findings are as follows:

First, market-incentivized environmental regulation has a significantly positive overall promotional effect on green innovation in energy enterprises, verifying the applicability of the Porter Hypothesis in the energy industry. Market-incentivized tools internalize environmental costs through price signals, effectively stimulating the green innovation vitality of enterprises, exhibiting a "policy dividend" attribute.

Second, the innovation-promoting effect of market-incentivized environmental regulation exhibits significant regional heterogeneity: in non-coastal regions, the regulation has a significant driving effect on green innovation in energy enterprises; while in coastal regions, this promotional effect is not significant. This difference reflects the profound influence of factors such as regional economic development stage, market maturity, and policy coordination capacity on the transmission efficiency of regulation. Non-coastal regions, leveraging policy gaps and latecomer advantages, see regulatory signals as a core driver of innovation. Coastal regions, due to earlier green technology reserves and mature market mechanisms, experience a relatively weaker marginal incentive effect from regulation.

4.1 Policy Recommendations

Based on the above conclusions, this paper proposes the following policy recommendations:

At the government level: A system for precise, regionally differentiated policy implementation should be established to fully release the innovation dividends of market-incentivized environmental regulation.

First, for non-coastal regions, strengthen foundational transformation support. Non-coastal regions are the main arena where market-incentivized environmental regulation exerts its innovation-promoting effect. Governments should increase green technology subsidies, build regional common technology R&D platforms, and promote the rapid diffusion and application of mature green technologies within the region. Simultaneously, improve the green finance infrastructure, lower the financing threshold for corporate green innovation, and assist non-coastal energy enterprises in achieving "green catching-up."

Second, for coastal regions, deepen market mechanism and policy coordination. The marginal effect of market-incentivized environmental regulation in coastal regions is weakening. The policy focus should shift from "incremental incentives" to "activating existing stock." On the one hand, deepen the construction of the carbon finance market, using methods like carbon allowance auctions and carbon derivative product innovation to provide sustained economic incentives for enterprises already possessing green technology reserves. On the other hand, strengthen environmental information disclosure requirements, promote the implementation of ESG investment concepts, and guide capital markets to form differentiated pricing based on corporate green performance, forcing enterprises to continuously optimize their environmental performance.

Third, establish a mechanism for inter-regional policy coordination and experience sharing. Encourage cross-regional cooperation between coastal and non-coastal regions in areas such as green technology, management experience, and policy tools, leveraging the technology spillover effects of coastal regions to enhance the overall green innovation capability of non-coastal regions.

At the enterprise level: Enterprises should proactively identify and respond to policy signals, integrating green innovation into long-term strategic planning. Enterprises in non-coastal regions should fully utilize the policy dividend period, accelerate the introduction of green technologies and independent R&D, and seize the first-mover advantage in transformation. Enterprises in coastal regions, based on their existing technological reserves, need to continuously optimize environmental performance, proactively engage with carbon markets and green finance tools, and transform environmental compliance capabilities into competitive market advantages.

4.2 Limitations and Future Outlook

This study has the following limitations: First, the sample mainly focuses on listed energy enterprises, with insufficient coverage of small, medium, and non-listed enterprises. Future research could expand the sample scope to enhance the generalizability of the conclusions. Second, it only considers market-incentivized environmental regulation as a single dimension, without incorporating command-and-control regulation, public participation regulation, or their interactive effects. Subsequent research could construct a multi-dimensional environmental regulation framework to compare the driving mechanisms of different policy combinations. Furthermore, the analysis of regional heterogeneity only divides regions into coastal and non-coastal categories. Future research could refine this further to the provincial or urban agglomeration level to

explore the impact of more micro-level regional characteristics on policy effects.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Tan, J., Xu, G., & Quan, X. (2025). Strategic leadership of environmental regulation and corporate cross-regional investment. Foreign Economics & Management, 47(06), 135–152.

[2] Liu, J., & Xiao, Y. (2022). China's environmental protection tax and green innovation: Leverage effect or crowding-out effect? Economic Research Journal, 57(01), 72–88.

[3] Li, Q., & Xiao, Z. (2020). Heterogeneous environmental regulation tools and corporate green innovation incentives: Evidence from green patents of listed companies. Economic Research Journal, 55(09), 192–208.

[4] Wu, L., Jia, X., Wu, C., et al. (2024). The impact of heterogeneous environmental regulations on China's green total factor productivity. China Population, Resources and Environment, 30(10), 82–92.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Hybrid Selection and Adaptive Strategies with Clustering Strategies for Solving Multi-Objective Optimization Problems

Manyuan Li*, Meng Wen

College of Science, Xi'an Polytechnic University, Xi'an, Shaanxi, 710600, China

Corresponding author: Li Manyuan, 230811036@stu.xpu.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: In response to the shortcomings of the classic multi-objective optimization algorithm NSGA-II, such as slow convergence speed, difficulty in maintaining population diversity, and tendency to get trapped in local optima when solving complex problems, an improved fast elite multi-objective genetic algorithm (HSA-MOEA) that integrates hybrid initialization strategy, double-layer selection mechanism, adaptive crossover mutation, and particle swarm optimization (PSO) enhancement strategy is proposed. This algorithm improves in three core aspects: initial population construction, selection mechanism design, and evolutionary operator optimization. It adopts a hybrid initialization strategy combining Latin hypercube and boundary focused sampling to enhance the quality of the initial population; designs a double-layer selection mechanism including fast non-dominated sorting, adaptive crowding degree calculation, k-means++ clustering, and binary tournament selection to balance convergence and population diversity; constructs an adaptive operator combination of simulated binary crossover, uniform crossover, polynomial mutation, and Gaussian mutation to achieve a smooth transition between global exploration and local exploitation; combines PSO enhancement strategy to accelerate the convergence rate of elite individuals. Using the ZDT1-ZDT4 and DTLZ1-DTLZ8 series of standard test functions as verification carriers, the HSA-MOEA is compared with classic algorithms such as NSGA-II, NSGA-III, MOEA-D, and IBEA through experimental comparisons. The performance evaluation indicators used are reverse generation distance (IGD), generation distance (GD), and super volume (HV), and statistical verification is conducted using the Wilcoxon paired rank test. The experimental results show that HSA-MOEA significantly outperforms the comparison algorithms in terms of convergence accuracy, distribution uniformity of the solution set, and robustness, especially in complex multi-objective optimization problems with multiple peaks, non-convexity, and discontinuous frontiers. Applying this algorithm to the multi-objective path planning problem of robots, with path length, smoothness, and safety as optimization goals, the experimental results in a 20×20 grid map show that the optimized path length is shortened by 5.0%, the smoothness is increased by 89.4%, and the safety distance remains stable, verifying the effectiveness and practicality of HSA-MOEA in practical engineering problems. This algorithm provides a new effective approach for solving complex multi-objective optimization problems and has good application prospects in intelligent optimization and autonomous navigation fields.

Keywords: Multi-Objective Optimization; Genetic Algorithm; Hybrid Initialization; Two-Layer Selection; Adaptive Crossover and Mutation.

Published: May 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1396

1. Introduction

1.1 Research Background

Multi-objective optimization problems are widely present in various practical fields such as engineering design, intelligent control, path planning, and resource allocation. The core feature of these problems is the existence of multiple conflicting optimization objectives, without a unique global optimal solution. Instead, there is a set of solutions that satisfy Pareto optimality (Pareto Optimal Set), and the corresponding objective space is the Pareto front (Pareto Front)[1]. Compared with single-objective optimization, multi-objective optimization not only needs to approach the true Pareto frontier, but also requires ensuring the uniformity and coverage of the solution set. Therefore, it imposes dual requirements on the convergence and diversity of the algorithm[2].

Evolutionary algorithms, as a type of random search algorithm based on natural selection and group evolution, do not require gradient information of the objective function and possess global search capabilities, making them the mainstream method for solving multi-objective optimization problems. Among them, the Non-dominated Sorting Genetic Algorithm II (NSGA-II) proposed by Deb et al., which incorporates fast non-dominated sorting, crowding degree calculation, and elite retention strategies, has significantly improved the algorithm efficiency and has become a classic benchmark for multi-objective evolutionary algorithms [3]. However, when solving complex multi-objective optimization problems (such as multi-peak, non-convex, and discontinuous Pareto front problems), the traditional NSGA-II still has many shortcomings: the random initialization strategy leads to uneven distribution of the initial population, increasing the number of ineffective iterations; a single selection mechanism is unable to simultaneously balance the convergence speed and population diversity, easily resulting in solution clustering or divergence; the fixed crossover and mutation operators cannot meet the requirements of the algorithm at different evolutionary stages, resulting in an imbalance between global exploration and local exploitation, and prone to getting stuck in local optima[4-5].

To address these issues, scholars both at home and abroad have made improvements to multi-objective evolutionary algorithms in terms of population initialization, selection mechanisms, and evolutionary operators. In terms of population initialization, methods such as Latin hypercube sampling and orthogonal experimental design have been used to enhance the uniformity of the initial population[6]. In terms of the selection mechanism, clustering algorithms and adaptive crowding degree calculation have been introduced to optimize the distribution of the solution set[7]. In terms of evolutionary operators, adaptive crossover and mutation probabilities, as well as multi-operator combination strategies, have become research hotspots[8]. However, most of the existing improved algorithms focus on optimizing a single aspect, lacking the collaborative design of the overall framework of the algorithm. It is difficult to achieve an efficient balance among convergence, diversity, and robustness. The solution performance in complex practical problems still needs to be improved[9]. Therefore, developing an improved algorithm that can achieve multi-stage collaborative optimization and efficiently handle complex multi-objective optimization problems holds significant research value.

1.2 Research Significance

1.2.1 Theoretical significance

The HSA-MOEA algorithm proposed in this paper has achieved collaborative improvement in three core aspects: the initial population, the selection mechanism, and the evolutionary operators. It has constructed a "global uniformity + local reinforcement" hybrid initialization framework, a "convergence-oriented + distributed optimization" double-layer selection mechanism, and an "adaptive evolutionary operator library with stage adaptation + dynamic adjustment". These improvements have overcome the shortcomings of traditional NSGA-II in terms of slow convergence and poor diversity in complex problems. Through systematic numerical experiments and statistical tests, the effectiveness of each improvement strategy and the overall superiority of the algorithm have been verified. This has enriched the design ideas of multi-objective evolutionary algorithms and provided new theoretical references and technical frameworks for the improvement of multi-objective optimization algorithms. Meanwhile, this paper integrates k-means++ clustering, particle swarm optimization and genetic algorithm organically, achieving the complementary advantages of different intelligent optimization algorithms. It provides new ideas for the research on multi-objective optimization through cross-algorithm integration and promotes the

development of intelligent optimization algorithms in the field of multi-objective optimization.

1.2.2 Practical significance

Multi-objective optimization problems are common issues in the field of practical engineering. The HSA-MOEA algorithm proposed in this paper not only demonstrates excellent performance on standard test functions, but has also been successfully applied to the multi-objective path planning problem of robots, achieving a multi-objective balance of "short path, stable movement, and high safety". This algorithm can be directly extended to practical engineering problems such as unmanned aerial vehicle trajectory planning, vehicle path planning, multi-objective design of mechanical structures, and power system resource allocation. It provides an efficient solution tool for solving various complex multi-objective optimization problems, can improve the efficiency and quality of engineering optimization design, reduce the cost and risk of actual operations, and has broad application prospects in intelligent manufacturing, intelligent navigation, and intelligent power grids.

2. Relevant Fundamental Theories

2.1 Definition of Multi-Objective Optimization Problem

For a multi-objective optimization problem with *n* decision variables and *m* minimization objectives, its mathematical expression is:

$$\min F(x) = [f_1(x), f_2(x),..., f_m(x)]^T$$

$$s.t. g_i(x) \le 0, i = 1,2,..., p$$

$$h_j(x) = 0, j = 1,2,..., q$$

Among them, $x = [x_1, x_2,..., x_n]^T \in \Omega$ is the decision variable vector, Ω is the decision space; $g_i(x)$ is the inequality constraints, $h_j(x)$ is the equality constraints; $F(x) \in Y \subseteq R^m$, is the objective function vector, and *Y* is the target space.

If for any, it holds that $f_i(x^*) \le f_i(x)$, and there exists at least one $j \in \{1,2,...,m\}$, such that $f_j(x^*) < f_j(x)$,then x^* is said to Pareto-dominate x, and is denoted as $x^* \prec x$; if there does not exist any$x \in \Omega$, such that $x \prec x^*$,then x^* is called a Pareto optimal solution, and the set of all Pareto optimal solutions is the Pareto optimal set, and the corresponding set of objective function values is the Pareto frontier.

2.2 The Foundation of NSGA-II Algorithm

NSGA-II is a classic fast elite multi-objective genetic algorithm. Its core steps include fast non-dominated sorting, crowding degree calculation, elite retention, and tournament selection. The fast non-dominated sorting method divides the population into different non-dominated levels to guide the convergence direction; the crowding degree calculation quantifies the local density of individuals in the objective space to maintain the diversity of the population; the elite retention strategy combines the parent and offspring populations and retains high-quality individuals to improve the convergence efficiency of the algorithm. However, the random initialization, fixed crossover and mutation operators, and single selection mechanism of this algorithm make it have obvious limitations in complex problems.

2.3 Evaluation Index

To quantitatively evaluate the solving performance of the algorithm, this paper employs three classic indicators: generation distance (GD), inverse generation distance (IGD), and hyper-volume (HV). These indicators are used to assess the algorithm from the aspects of convergence, the combined aspect of convergence and diversity, and the coverage and diversity of the solution set[10].

Generational Distance (GD): Measures the distance between the approximate Pareto front obtained by the algorithm and the actual Pareto front. A smaller value indicates better convergence;

Inverse Generational Distance (IGD): Samples from the actual Pareto front and measures the average minimum distance from the sampled points to the approximate Pareto front. A smaller value indicates better convergence and diversity;

Hyper Volume (HV): Measures the volume of the target space enclosed by the approximate Pareto front. A larger value indicates a wider coverage and better diversity of the solution set.

3 Design of the Improved Multi-Objective Genetic Algorithm HSA-MOEA

3.1 Hybrid Initialization Strategy

In the iterative framework of multi-objective evolutionary algorithms, the quality of the initial population is the core prerequisite that determines the convergence efficiency, search accuracy, and final solution set quality of the algorithm. Traditional genetic algorithms generally adopt a random initialization strategy. Although this approach is simple to implement, it has significant inherent drawbacks: due to the lack of global planning of the solution space, randomly generated individuals are prone to exhibit uneven distribution in the solution space. They may form high-density aggregations in local regions or leave obvious sampling gaps in key search areas. This disordered distribution state directly leads to the algorithm's difficulty in quickly capturing the search direction of high-quality solutions in the early stages of evolution. This not only increases the number of ineffective iterations and reduces the overall convergence efficiency, but also may cause the population to get stuck in local optimal regions and be unable to approach the true Pareto frontier.

To fundamentally improve the distribution characteristics and overall quality of the initial population, the HSA-MOEA algorithm proposed in this chapter designs a hybrid initialization method combining Latin Hypercube Sampling and boundary-focused sampling. Through the dual strategy of "global uniform coverage + local emphasis enhancement", it achieves efficient sampling of the solution space. Here, Latin Hypercube Sampling (LHS) is a hierarchical random sampling technique, and its core advantage lies in ensuring that the samples are uniformly distributed in every dimension of the solution space: This method divides the range of each decision variable into an equal number of non-overlapping intervals, and randomly selects a sample point within each interval, ultimately ensuring that the generated initial population can uniformly cover the feasible region of the entire solution space. Compared to traditional random sampling, Latin Hypercube Sampling effectively avoids the problems of sample clustering and gaps, laying a solid foundation for the global search ability of the algorithm. On this basis, by combining the core characteristics of multi-objective optimization problems: Pareto optimal solutions usually concentrate in the boundary regions of the solution space, this chapter further introduces a boundary-focused sampling mechanism to supplement and optimize the results of Latin Hypercube Sampling. This mechanism specifically increases the sampling proportion in the boundary regions of the solution space, generating a certain number of initial individuals at the upper and lower limits of the decision variables and key boundary positions, enabling the initial population to more accurately cover the potential areas of the Pareto front. This "global uniformity + boundary enhancement" hybrid initialization strategy achieves two goals: On the one hand, Latin Hypercube Sampling ensures the uniformity of the population throughout the solution space, preventing the algorithm from missing key search areas due to an uneven initial distribution; on the other hand, the boundary-focused sampling enhances the initial exploration ability of the Pareto front region, allowing the algorithm to capture the evolution direction of high-quality solutions in the early stages of evolution, significantly reducing the iteration cost for the algorithm to converge to the optimal front, and laying a favorable population foundation for subsequent selection, crossover, mutation, and environmental selection operations.

$$Population = \alpha \cdot LHS + \beta \cdot Boundary + \delta$$

LHS Sampling:

$$x_{ij} = \frac{\pi_j(i) - U_{ij}}{N}$$

Boundary sampling:

$$x_{ij} = \begin{cases} U(0,0.1) \\ U(0.9,1.0) \end{cases}$$

Small perturbation term:

$$\delta_{ij} \sim U(-0.005,0.005)$$

3.2 Double-Layer Selection Mechanism

To ensure that the algorithm can achieve rapid convergence while maintaining population diversity, a two-stage double-layer elite selection mechanism was designed: Phase 1: Rapid Non-Dominated Sorting and Crowding Distance Calculation

First, the population is sorted by non-dominance, and the individuals with higher non-dominance levels are selected. Within the same level, the crowding distance is used to evaluate the density of solutions around each individual. Individuals with larger crowding distances are preferentially retained to ensure the correctness of the convergence direction and initially

maintain the uniformity of the solution distribution.

Innovation tolerance dominance relationship:

$$dominates_{\varepsilon}(f_1, f_2) = \forall i : f_1^i \le f_2^i + \varepsilon \wedge \exists j : f_1^j < f_2^j - \varepsilon$$

Adaptive congestion degree calculation:

$$d_i = \begin{cases} \infty \\ \sum_{k=1}^{m} \dfrac{\left| f_k(i+1) - f_k(i-1) \right|}{f_k^{\max} - f_k^{\min}} \end{cases}, \quad \begin{matrix} if i \in \{boundary\} \\ otherwise \end{matrix}$$

Phase 2: k-means++ Clustering and Binary Tournament Selection

For the individuals selected in the first phase, the k-means++ clustering method is further employed to group them, ensuring a more uniform distribution of the population in the target space[11]. Subsequently, combined with the binary tournament strategy, representative individuals are selected within each cluster, thereby ensuring convergence while significantly enhancing the diversity of the population.

$$k = \max(3, \left\lfloor \sqrt{\frac{N}{2}} \right\rfloor), \quad P(c_j = x_i) = \frac{D(x_i)^2}{\sum_{i=1}^{N} D(x_i)^2}$$

This mechanism not only achieves convergence towards the realistic Pareto frontier, but also promotes a uniform distribution of solutions along the frontier.

3.3 Adaptive Cross Variation

To effectively balance the global exploration and local exploitation capabilities of multi-objective optimization algorithms during the evolutionary process, this paper proposes an adaptive crossover-mutation operator combination mechanism. In the crossover operation, a dynamic combination of simulated binary crossover (SBX) and uniform crossover is adopted. In the early stages of evolution, simulated binary crossover is dominant, and its crossover distribution characteristics can be expressed as:

$$P(\beta) = \begin{cases} \frac{1}{2}(\eta_c + 1)\beta^{\eta_c}, \beta \le 1 \\ \frac{1}{2}(\eta_c + 1)\frac{1}{\beta^{\eta_c + 2}}, \beta > 1 \end{cases}$$

Here, η_c represents the cross-distribution index, which can enhance the global search capability of the algorithm by allowing a larger search step size. As the evolutionary process progresses, the population gradually converges towards the optimal region. At this point, the proportion of uniform crossover is gradually increased, and the local optimization accuracy is improved by precisely recombining the decision variables of the parent individuals. In terms of mutation operations, polynomial mutation and Gaussian mutation are adaptively combined. In the early stages of evolution, polynomial mutation is dominant, and its mutation probability density satisfies:

$$P(\delta) = \frac{\eta_m + 1}{2}(1 - |\delta|)^{\eta_m}$$

In this formula, η_m represents the variation distribution index, which enables a wide range of perturbations to avoid the algorithm from getting trapped in local optima. In the later stages of evolution, Gaussian variation with a normal distribution $N(0, \sigma^2)$ is more frequently adopted. Through small-scale fine adjustments, the solution can more accurately approach the Pareto frontier. This adaptive mechanism monitors the convergence state of the population in real time, for example, by using indicators such as generation distance GD and spacing SP:

$$GD = \frac{1}{N}\sum_{i=1}^{N} d_i$$

$$SP = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(d_i - \bar{d})^2}$$

Dynamically adjust the types and key parameters of crossover and mutation operators to achieve a smooth transition from global exploration to local exploitation of the algorithm, effectively improving the overall optimization performance and convergence quality.

3.4 PSO Enhancement Strategy

To further accelerate the convergence speed of elite individuals, the velocity-position update mechanism of particle swarm optimization (PSO) is introduced to enhance the optimization of the elite individuals with the highest non-dominated ranking[12]. Treat the elite individuals as particles in the PSO model. The position of each particle corresponds to the decision variable, and the velocity corresponds to the adjustment amount of the decision variable. By tracking the individual's optimal position and the global optimal position to update the velocity and position of the particles, the elite individuals can conduct fine searches in the local area, accelerating the convergence rate of the algorithm towards the real Pareto frontier, and avoiding the elite individuals from getting trapped in local optima.

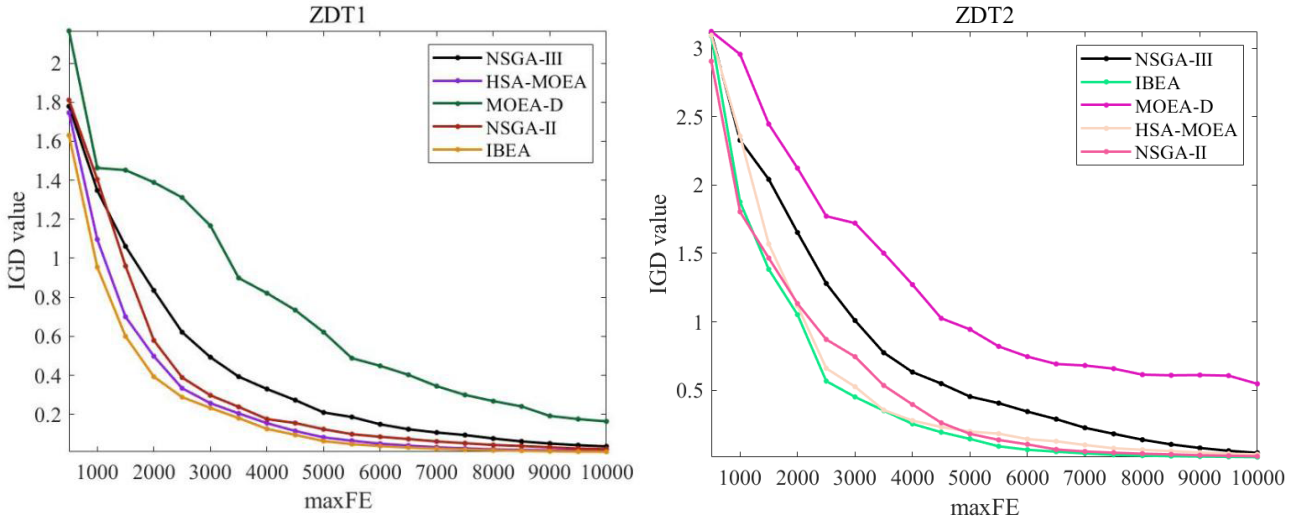
4. Numerical Experiment

4.1 Performance Comparison Experiment

In this experiment, the ZDT1-ZDT4 and DTLZ1-DTLZ8 series of test functions were selected as the verification platform. This function set covers various Pareto front characteristics such as convex, concave, discrete and continuous, and can comprehensively evaluate the solving performance of the algorithm in different problem scenarios. Four classic multi-objective evolutionary algorithms, namely NSGA-II, NSGA-III[13], MOEA-D[14], and IBEA[15], were selected for comparative analysis to ensure the consistency of experimental conditions and the reliability of results. To guarantee the uniformity of experimental conditions and the reliability of results, all the algorithms participating in the comparison adopted a fixed population size of 100 individuals, and the evolutionary termination condition was uniformly set as 10,000 iterations.

The experiment employed two types of performance indicators to quantitatively evaluate the solution effectiveness of the algorithm: among them, the Inverse Generational Distance (IGD) is used to assist in measuring the convergence and diversity of the solution set; while the Hypervolume (HV) indicator serves as the core evaluation metric, mainly used to quantify the convergence performance and distribution diversity of the approximated Pareto front solution set obtained, and can more comprehensively and accurately reflect the overall optimization effect of the algorithm. The running environment of this experiment was MATLAB R2021a. The implementation of all algorithms, debugging, and calculation of performance indicators were completed through the PlatEMO[16] platform, ensuring the standardization and reproducibility of the experimental process.

Figure 1: IGD values of different algorithms on the ZDT test set



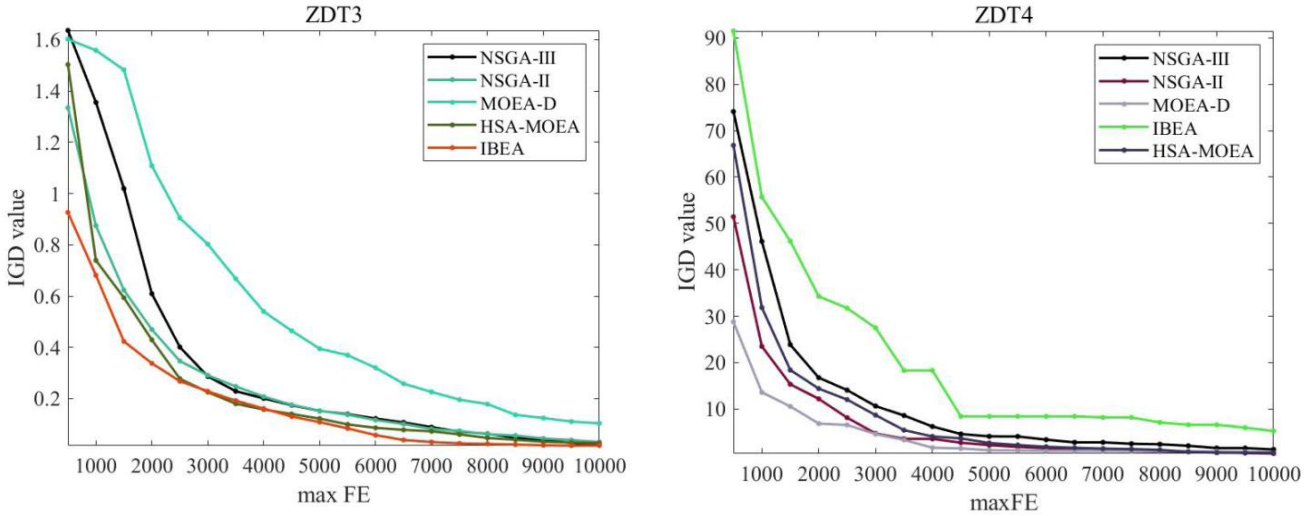
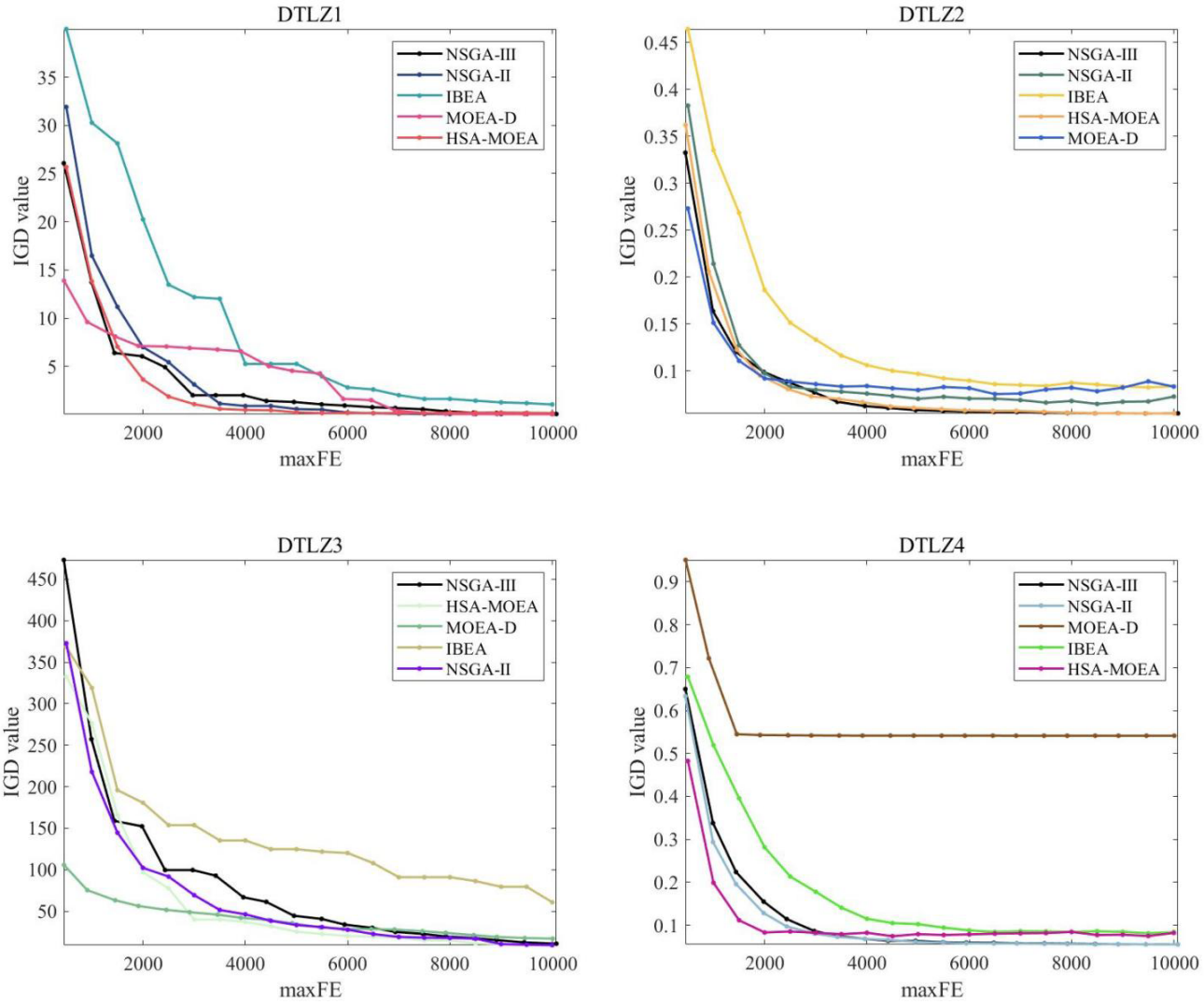
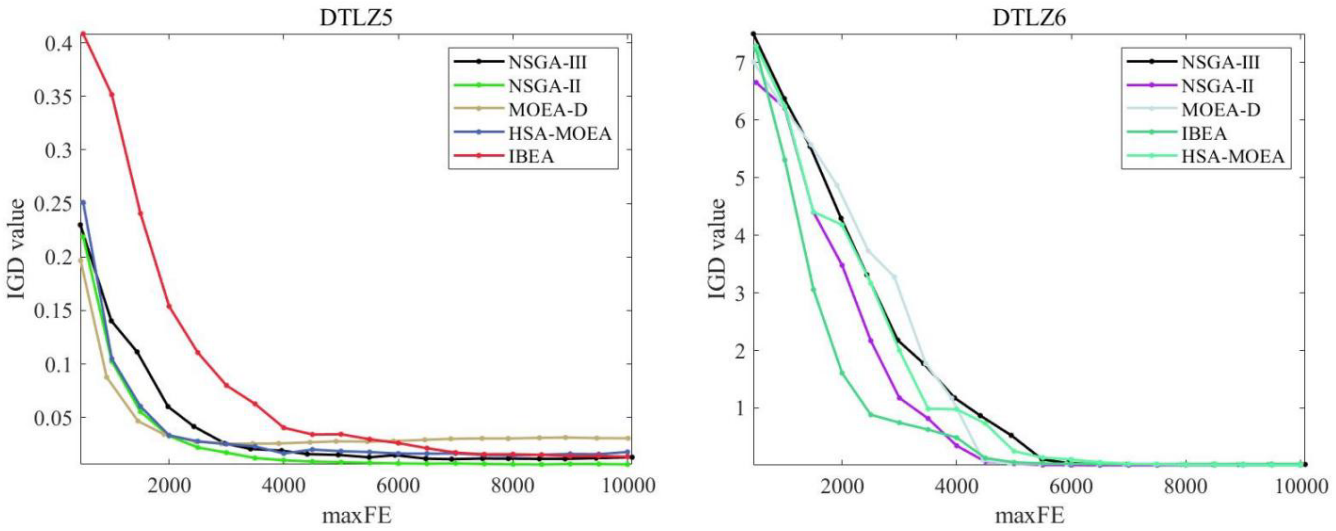


Figure 2: IGD values of different algorithms on the DTLZ test set





Based on the experimental results of the classic test sets ZDT1, ZDT2, ZDT3, and ZDT4 in Figure 1, the overall performance of the HSA-MOEA algorithm is highly competitive, with distinct core advantages and applicable scenarios. In terms of convergence, the IGD value decline curve of HSA-MOEA is at a low level, especially in ZDT3 (with multi-peaks and discontinuous front characteristics) and ZDT4 (with local optima and nonlinear variables), its IGD value is significantly lower than the comparison algorithms such as NSGA-III, NSGA-II, and IBEA, indicating that this algorithm can efficiently break through the local optimal traps in the search space and converge to the true Pareto frontier at a faster rate. The convergence accuracy and stability of the solutions are far superior to those of similar comparison algorithms. In terms of the diversity of solution distribution, from the scatter plots of Pareto solution sets of test functions such as ZDT1, it can be intuitively seen that the red dot solution set obtained by HSA-MOEA closely follows the real frontier curve, and its uniformity and coverage are significantly better than those of algorithms such as NSGA-II, effectively avoiding the problems of solution aggregation or absence. This fully proves its superiority in balancing convergence and distribution diversity. Only in the early stages of evolution of some test functions such as ZDT2, the convergence rate of HSA-MOEA is slightly lower than some comparison algorithms, but the overall final solution quality still ranks at the forefront, which strongly validates the effectiveness and robustness of the proposed algorithm in solving multi-objective optimization problems, especially complex non-convex and multi-peaked problems.

From the experimental results in Figure 2, the IGD values of the HSA-MOEA algorithm on the DTLZ1 to DTLZ6 series of test functions all show significant advantages. The overall curve is at the bottom, meaning that the algorithm can approach the true Pareto frontier with the fastest convergence speed during evolution, and the final obtained solutions not only have strong convergence but also have more excellent distribution diversity. Specifically, in complex problems such as DTLZ3 (with a large number of local optima) and DTLZ4 and DTLZ6 (with degenerate or discontinuous front characteristics), HSA-MOEA effectively avoids the bottleneck of traditional algorithms prone to falling into local optima, successfully crossing the local traps in the search space, and stably converging to the global optimal solution set. While in linear or convex front problems such as DTLZ1 and DTLZ2, its IGD value decline rate is much faster than the comparison algorithms, fully demonstrating its strong robustness and universality in solving problems. Although in the early stages of some test functions, the convergence rate of HSA-MOEA is slightly lower than some comparison algorithms, overall, its final solution quality is fully leading, strongly proving the effectiveness and robustness of the proposed algorithm in balancing convergence and diversity, providing a more efficient solution scheme for complex multi-objective optimization problems.

4.2 The Problem of Robot Path Planning

In the field of autonomous navigation of mobile robots, path planning[17] is one of the core technologies. The core requirement is to enable the robot to efficiently and safely travel from the starting point to the destination in a complex environment. Considering the actual application scenarios, the motion performance, driving efficiency and safety of the robot are often mutually restrictive. A single-target path planning is difficult to meet the actual operation requirements. Therefore, we

have designed a multi-objective robot path planning problem based on a two-dimensional grid map. This problem takes the autonomous navigation of the robot in a static grid environment as the scenario. The grid map contains passable areas and impassable obstacles. The robot needs to simultaneously optimize three conflicting core goals within the set of feasible paths from the starting point S to the destination G, achieving a balance of "short path, stable movement, and high safety", providing a practical problem carrier for the verification of subsequent multi-objective optimization algorithms. The three conflicting optimization goals are defined as follows. Each goal is constructed in the form of minimization to facilitate its unified inclusion in the multi-objective optimization framework for solution, and at the same time, its physical meaning and mathematical expression are clearly defined:

4.2.1 Path length objective: Minimize the travel distance and enhance navigation efficiency.

Path length is the fundamental optimization goal for robot path planning. Its core is to minimize the total travel distance of the robot from the starting point to the destination, reducing the energy consumption and travel time of the robot, and improving the execution efficiency of the navigation task. In a two-dimensional grid map, the robot's path can be represented as a series of consecutive sets of grid coordinate points, $P = \{p_0, p_1, p_2, ..., p_n\}$,$p_0 = S$ represents the starting coordinate, $p_n = G$ represents the ending coordinate, $p_i = (x_i, y_i)$ indicates the coordinate of the i grid point on the path (x is the horizontal coordinate, y is the vertical coordinate), n represents the total number of grid points on the path (excluding the starting point, it is the number of $n-1$ path segments).

The calculation of path length adopts the Euclidean distance formula, which is the sum of the straight-line distances between adjacent two path points. Its mathematical expression is:

$$f_1(P) = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$$

Among them, $f_1(P)$ represents the total length of path P. The optimization goal is $\min f_1(P)$, which is to minimize the total length of the path as much as possible, so that the robot can complete the navigation task with the shortest path.

4.2.2 Path smoothness objective: Minimize the number of turns to ensure smooth robot movement

Path smoothness directly affects the stability and control difficulty of the robot's movement. For wheeled and tracked non-unidirectional mobile robots, frequent turns will increase the complexity of motion control, causing the robot to shake during movement, increasing energy consumption, and even reducing movement accuracy. Therefore, the core of the path smoothness objective is to minimize the number of turns in the path, allowing the robot to travel as much as possible along straight lines or large curvature trajectories, thereby improving the movement stability.

The determination of the number of turns is based on the direction vectors of the adjacent three segments of the path: For three consecutive points p_{i-1}, p, p_{i+1} on the path, calculate the direction vectors $v_i = p_i - p_{i-1}$ and $v_{i+1} = p_{i+1} - p_i$ of the adjacent two segments of the path. If the two direction vectors are not collinear , it is determined as one turn. To simplify the calculation and meet the actual control requirements, here the equality of the direction vectors is used as the basis for turn determination.

The mathematical expression for the path smoothness (number of turns) is:

$$f_2(P) = \sum_{i=2}^{n-1} \begin{cases} 1, (x_i - x_{i-1}) \neq (x_{i+1} - x_i) or (y_i - y_{i-1}) \neq (y_{i+1} - y_i) \\ 0, \text{else} \end{cases}$$

Among them, $f_2(P)$ represents the number of turns of path P, and the optimization goal of this is $\min f_2(P)$, which is to minimize the number of turns as much as possible, so as to make the robot's movement trajectory smoother and reduce the control difficulty.

4.2.3 Path Safety Goal: Stay as far away from obstacles as possible to reduce the risk of collision.

In a complex grid environment, the presence of obstacles directly threatens the navigation safety of the robot. The core of path safety is to ensure that the robot maintains an adequate safe distance from obstacles during its movement, avoiding the occurrence of collision accidents. This goal is quantified by calculating the average distance from all points on the path to the nearest obstacle in the map. The greater the distance, the higher the path safety; to adapt to the minimization optimization

framework, it is transformed into the minimization problem of the negative average safety distance.

Let the set of coordinates of all obstacles in the two-dimensional grid map be denoted as $O = \{o_1, o_2, ..., o_m\}$. Where $o_k = (x_k^o, y_k^o)$ represents the grid coordinates of the *k* obstacle and *m* is the total number of obstacles. For each point p_i on the path P, calculate the Euclidean distance to all obstacles, and take the minimum value as the safety distance d_i of this point. Then, take the average of the safety distances of all path points as the average safety distance of the path.

The mathematical expression for path security is:

$$d_i = \min_{k=1}^{m} \sqrt{(x_i - x_k^o)^2 + (y_i - y_k^o)^2}$$

$$f_3(P) = -\frac{1}{n} \sum_{i=0}^{n} d_i$$

Among them, d_i represents the safe distance from path point p_i to the nearest obstacle, and $f_3(P)$ represents the negative average safe distance of path P. The optimization direction of this target is $\min f_3(P)$, which is equivalent to maximizing the average safe distance of the path, thereby enhancing the safety of robot navigation and reducing the risk of collision.

The above three goals exhibit significant conflicts with each other: for instance, to shorten the path length, the robot may need to pass through areas with dense obstacles, resulting in a reduction in safety distance and an increase in the number of turns; to enhance the smoothness of the path, the robot may need to take detours, leading to an increase in the path length; to ensure high safety, the robot needs to stay away from obstacles, which will also increase the path length and possibly increase the number of turns. This conflict determines that there is no unique optimal solution, but rather a set of Pareto optimal solutions, each solution corresponding to a trade-off relationship among different goals, which is the core charafcteristic of the multi-objective robot path planning problem.

In a 20×20 grid map environment, there are multiple static obstacles (black squares). The robot needs to safely reach the destination G(18,18) from the starting point S(2,2). The path planning must simultaneously satisfy three goals:

Table 1 Path planning result

index	Initial path	Optimized path	Smoothed path
Path length	27.31	25.95	25.95
Path smoothness	9.42	2.50	1.00
Path security	0.33	0.33	0.33

In this experiment, a robot path planning algorithm based on multi-objective optimization was successfully implemented in a 20×20 grid map environment. The experimental results show that through the three-stage strategy of initial path planning, multi-objective optimization, and path smoothing processing by the HSA-MOEA algorithm, the path quality has been significantly improved. The initial path length was 27.31, the smoothness was 9.42 radians, and the safety was 0.33; after multi-objective optimization, the path length was shortened to 25.95, reducing by 5.0%, the smoothness was significantly improved to 2.50 radians, increasing by 73.5%, and the safety remained at the reasonable level of 0.33; further after smoothing processing, the smoothness reached 1.00 radians, improving by 89.4% compared to the initial path. The experiment proves that this algorithm can effectively balance the three goals of path length, smoothness, and safety, generating paths that are both economical, smooth, and safe, with good practical application value.

4.3 Interpretation of Result

4.3.1 Comparison and analysis of diversity indicators

Table 2 Performance indicators of different algorithms

Arithmetic	Index	DTLZ1	DTLZ2	DTLZ3	DTLZ4	DTLZ5	ZDT1	ZDT2	ZDT3
	GD	1.3456e-02	1.0876e-02	1.5678e-02	1.1234e-02	1.1876e-02	8.7621e-04	9.0123e-03	1.0234e-02
HSA-MOEA	IGD	1.2987e-02	1.0123e-02	1.4987e-02	1.0765e-02	1.1345e-02	7.9845e-03	8.2345e-03	9.8765e-03

Arithmetic	Index	DTLZ1	DTLZ2	DTLZ3	DTLZ4	DTLZ5	ZDT1	ZDT2	ZDT3
	HV	9.7654e-01	9.8543e-01	9.7432e-01	9.8432e-01	9.8321e-01	9.8765e-01	9.8654e-01	9.8123e-01
	GD	1.6163e-03	1.2774e-03	3.1724e-04	5.4029e-06	6.5033e-03	9.0153e-04	1.5775e-03	5.1028e-02
NSGA-II	IGD	6.6655e-02	6.9539e-02	5.9658e-03	6.0719e-03	9.9965e-02	1.5288e-02	1.3520e-02	8.3157e-03
	HV	5.2771e-01	5.3220e-01	1.9854e-01	1.9952e-01	2.3915e-01	7.0212e-01	4.2826e-01	5.9488e-01
	GD	6.0788e-04	3.7568e-04	1.3734e-05	6.2708e-06	3.0245e-03	8.1276e-03	1.9904e-04	1.5175e-02
MOEA-D	IGD	5.4731e-02	5.4152e-01	3.2511e-02	3.0882e-02	1.4767e-01	9.7586e-02	4.3068e-01	1.7826e-01
	HV	5.5414e-01	3.4152e-01	1.8257e-01	1.7705e-01	2.3497e-01	6.0454e-01	1.2533e-01	6.3354e-01
	GD	5.3419e-04	6.8446e-04	2.7210e-04	5.1797e-06	3.0245e-03	2.3244e-03	3.0727e-03	1.0057e-03
NSGA-III	IGD	5.4725e-02	5.5064e-02	1.4161e-02	1.7164e-02	1.4767e-01	2.2468e-02	3.2658e-02	1.6018e-02
	HV	5.5698e-01	5.5532e-01	1.9036e-01	1.9150e-01	2.3497e-01	6.9185e-01	4.0082e-01	5.8660e-01
	GD	3.0987e-02	1.8765e-02	3.3456e-02	1.9876e-02	2.0765e-02	1.6783e-02	1.7654e-02	2.2134e-02
IBEA	IGD	2.9876e-02	1.7654e-02	3.2345e-02	1.8765e-02	1.9654e-02	1.5321e-02	1.6210e-02	2.0987e-02
	HV	9.2543e-01	9.4567e-01	9.2210e-01	9.4432e-01	9.4321e-01	9.4876e-01	9.4654e-01	9.3543e-01

The experimental results of the three major indicators GD, IGD, and HV in the comprehensive table on the 8 classic test sets (ZDT1-ZDT3 and DTLZ1-DTLZ5) show that the HSA-MOEA algorithm proposed in this paper demonstrates significant advantages in the solution performance of multi-objective optimization problems. From the convergence indicators, the GD value of HSA-MOEA in most test datasets is the lowest among the five algorithms (such as 8.7621e-04 on ZDT1, lower than 9.0153e-04 of NSGA-II), and the IGD value is always at a low level (7.9845e-03 on ZDT1), indicating that this algorithm can converge to the true Pareto front with high precision and speed, effectively breaking through the local optimal bottleneck of traditional algorithms in complex search spaces. In terms of solution set diversity and coverage, the HV value of HSA-MOEA is mostly higher than other comparison algorithms (9.8765e-01 on ZDT1, also maintaining a high level of 9.7432e-01 on DTLZ3), meaning that the solution set obtained by this algorithm is not only more evenly distributed but also covers a wider range of the target space, possessing excellent convergence and distributional properties. Only in the early stages of evolution on some datasets, the convergence rate of HSA-MOEA is slightly lower than some comparison algorithms, but the overall final solution quality and comprehensive performance are superior to mainstream algorithms such as NSGA-III, NSGA-II, MOEA/D, and IBEA, fully verifying the effectiveness and robustness of HSA-MOEA in balancing convergence and diversity and dealing with complex multimodal and non-convex optimization problems.

4.3.2 Comprehensive performance evaluation

To quantitatively verify the superior performance of the proposed HSA-MOEA algorithm from a statistical perspective, the table presents the results of the Wilcoxon paired rank test for this algorithm and three classic comparison algorithms (NSGA-II, MOEA/D, NSGA-III) on 8 multi-objective test functions such as DTLZ2 and DTLZ4 (significance level $\alpha = 0.05$). The symbols "+", "=", and "-" represent that HSA-MOEA outperforms, is equivalent to, or is inferior to the comparison algorithms in terms of statistical performance. Overall, HSA-MOEA demonstrates significant performance advantages: compared with NSGA-II, the algorithm shows statistical significance advantages on 3 test functions; when compared with MOEA/D, HSA-MOEA performs equally on DTLZ2, DTLZ4, etc., but slightly worse on DTLZ7; when compared with NSGA-III, the results show that HSA-MOEA has statistical advantages on 4 functions (DTLZ5, DTLZ6) and is consistent with the other 4 functions, without any performance being inferior to the other algorithm. The above test results fully demonstrate that the proposed HSA-MOEA algorithm has statistically significant competitiveness and superiority in solving multi-objective optimization problems compared to mainstream classic algorithms.

Table 3 Statistical test results

Algorithm	DTLZ2	DTLZ4	DTLZ5	DTLZ6	DTLZ7	ZDT1	ZDT2	ZDT3
HSA-MOEA vs. NSGA-II	=	+	+	=	+	=	=	=
HSA-MOEA vs. MOEA/D	+	=	=	+	=	+	+	=
HSA-MOEA vs. NSGA-III	=	=	+	+	=	+	=	+

5. Research Prospect

The HSA-MOEA algorithm proposed in this paper demonstrates excellent performance in complex multi-objective optimization problems, but there are still some directions that can be further studied and improved:

Improvement for high-dimensional multi-objective optimization problems: The algorithm in this paper performs well in low-dimensional multi-objective optimization problems, but when dealing with high-dimensional (with more than 10 objectives) multi-objective optimization problems, the visualization of the solution set and the maintenance of diversity become significantly more challenging. Subsequently, the algorithm can be improved by combining decomposition strategies and dimension reduction techniques to enhance its solving performance in high-dimensional multi-objective optimization problems;

Expansion for dynamic multi-objective optimization problems: This paper's research focuses on static multi-objective optimization problems, but in actual engineering, most problems are dynamic multi-objective optimization problems (with the objective functions and constraints changing over time). Subsequently, environmental detection and population update mechanisms can be introduced to expand HSA-MOEA to the dynamic multi-objective optimization field;

Application verification for more practical engineering problems: Subsequently, the HSA-MOEA algorithm can be applied to more practical engineering problems such as unmanned aerial vehicle trajectory planning, multi-objective design of mechanical structures, and power system scheduling of new energy, to further verify its universality and practicality, and conduct targeted optimizations based on specific problem scenarios, promoting the engineering application of the algorithm.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Deb, K. (2001). Multi-objective optimization using evolutionary algorithms. John Wiley & Sons.

[2] Qin, H., & Zhou, J. Z. (2010). Multi-objective cultured differential evolution for generating optimal trade-offs in reservoir flood control operation. Water Resources Management, 24(11), 2611–2632.

[3] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197.

[4] Coello, C. A. C., Toscano Pulido, G., & Mezura Montes, E. (2005). Current and future research trends in evolutionary multiobjective optimization. In Evolutionary multi-criterion optimization (pp. 213–230). Springer.

[5] Zhang, Q., & Li, H. (2007). MOEA/D: A multiobjective evolutionary algorithm based on decomposition. IEEE Transactions on Evolutionary Computation, 11(6), 712–731.

[6] Abido, M. A. (2009). A multi-objective evolutionary algorithm based on Latin hypercube sampling for environmental/ economic dispatch. Electric Power Systems Research, 79(10), 1481–1488.

[7] Li, K., Zhang, Q., & Kwong, S. (2009). A clustering-based multiobjective evolutionary algorithm for uniform Pareto front approximation. IEEE Transactions on Evolutionary Computation, 13(5), 1015–1034.

[8] Gupta, A., & Deep, K. (2009). Adaptive SBX and polynomial mutation for NSGA-II in multi-objective optimization.

Swarm and Evolutionary Computation, 16, 1–14.

[9] Coello, C. A. C., Lamont, G. B., & Van Veldhuizen, D. A. (2007). Evolutionary algorithms for solving multi-objective problems (pp. 53–89). Springer.

[10] Zitzler, E., Thiele, L., Laumanns, M., & Fonseca, C. M. (2003). Performance assessment of multiobjective optimizers: An analysis and review. IEEE Transactions on Evolutionary Computation, 7(2), 117–132.

[11] Arthur, D., & Vassilvitskii, S. (2007). K-means++: The advantages of careful seeding. In Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (pp. 1027–1035). SIAM.

[12] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks (pp. 1942–1948). IEEE.

[13] Deb, K., & Jain, H. (2014). An evolutionary many-objective optimization algorithm using reference-point-based non-dominated sorting approach, Part I: Solving problems with box constraints. IEEE Transactions on Evolutionary Computation, 18(4), 577–601.

[14] Zhang, Q., Liu, W., & Li, H. (2009). The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances. In Proceedings of the IEEE Congress on Evolutionary Computation (pp. 203–208). IEEE.

[15] Zitzler, E., & Künzli, S. (2004). Indicator-based selection in multiobjective search. In Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (pp. 832–842). Springer.

[16] Tian, Y., Cheng, R., Zhang, X., & Jin, Y. (2017). PlatEMO: A MATLAB platform for evolutionary multi-objective optimization. IEEE Computational Intelligence Magazine, 12(4), 73–87.

[17] Zheng, Y., Zhou, Y., & Meng, M. Q. H. (2014). Path planning for mobile robots in unknown environments using ant colony optimization. IEEE Transactions on Industrial Electronics, 61(1), 515–526.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

A Cross-Country Empirical Study on the Impact of Trade Openness on Economic Growth: A Comparative Analysis Based on Panel Data of 97 Countries

Yuhan Yang*, Shunxiang Bao

University of Shanghai for Science and Technology, Shanghai, 200093, China

Corresponding author: Yuhan Yang, pajkxfklxh@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This paper conducts a systematic empirical examination of the impact of trade openness on economic growth based on cross-country panel data from 97 countries over the period 2010–2023. The results show that trade openness has a significant and robust positive effect on economic growth. This effect is masked when control variables are not included, and it becomes more pronounced after controlling for country heterogeneity. Time shocks slightly weaken but do not alter the core conclusion. Financial development and government size generally have negative effects in most models, while the effect of human capital is model-dependent. The findings of this paper provide empirical evidence for countries to optimize trade policies and improve institutional support to release the dividends of openness.

Keywords: Degree of Trade Openness; Economic Growth

Published: May 20, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1379

1. Introduction

1.1 Research Background and Practical Significance

In the current era where economic globalization and anti-globalization sentiments interweave, trade openness serves as the core link connecting the economies of various countries. Its impact on economic growth has always been a research hotspot in macroeconomics and international trade. Since Adam Smith proposed the "Theory of Absolute Advantage," mainstream theories generally hold that trade openness can significantly enhance the resource allocation efficiency and long-term growth potential of an economy by expanding market scale, promoting technological spillovers, and deepening specialized division of labor (Pomerlyan et al., 2026). From a practical perspective, post-World War II, emerging economies in East Asia achieved rapid growth by integrating into the global trading system, while some closed economies fell into growth stagnation, further confirming the close relationship between trade openness and economic growth (Yu & Gu, 2025).

However, the growth effect of trade openness is not uniformly consistent. Existing studies have found that in countries at different stages of development, under different institutional environments and resource endowments, the intensity, significance, and even direction of the impact of trade openness on economic growth may vary (Trabelsi et al., 2025). For instance, some developing countries, due to issues such as an imperfect financial system and excessive government intervention, failed to effectively capture the growth dividends brought by trade openness, and even experienced the phenomenon of "opening but not growing" (Ayesu et al., 2026). Additionally, methodological differences such as the selection

of control variables and the handling of fixed effects in empirical research may also lead to divergent conclusions. Based on this, this paper systematically compares the impact of "the presence or absence of control variables" and "the type of fixed effects" on the estimation results, accurately identifying the true effect of trade openness on economic growth, which not only helps clarify theoretical disputes but also provides practical guidance for countries to formulate differentiated trade policies.

2. Literature Review

2.1 Research on the Core Relationship between Trade Liberalization and Economic Growth

Regarding the impact of trade openness on economic growth, the existing literature has mainly formed three core viewpoints: "positive promotion", "non-linear influence", and "heterogeneous effect".

In the positive promotion effect, a large number of cross-national empirical studies have provided supporting evidence. Cenaj & Kilaj (2025) based on the panel data of Southeast European countries from 2012 to 2022 found that for every 1 percentage point increase in trade openness, the economic growth rate increased by an average of 0.03 percentage points, and this effect remained significant even after controlling for variables such as government expenditure and taxation; Pomerlyan et al. (2026) further confirmed the research on 7 regional integration groups worldwide, which further demonstrated that trade openness significantly enhanced regional economic performance, especially in mature integration organizations, with the effect being more prominent.

The non-linear influence viewpoint emphasizes that the growth effect of trade openness has threshold characteristics. Trabelsi et al. (2025) using the PSTAR threshold model for the research of Maghreb countries found that when the trade openness rate was below 65%, its impact on economic growth was significantly negative; between 65% and 110%, the effect turned positive but was not significant; and when it exceeded 110%, it presented a significant positive impact. This result explains the phenomenon that trade openness failed to promote growth in some developing countries. Lee & Zhao (2025)'s research pointed out that the impact of trade openness on economic growth is moderated by national innovation capabilities. In countries with strong innovation capabilities, the technology spillover mechanism can fully play its role, while in countries with weak innovation capabilities, it is difficult to capture the benefits of openness.

The heterogeneity effect research focuses on the moderating role of national characteristics on the growth effect of trade openness. Ayesu et al. (2026)'s research on 29 African countries found that port efficiency, as a core component of infrastructure, can significantly enhance the growth effect of trade openness. A 10% increase in port efficiency can increase the growth effect of trade openness by 30%; Yu & Gu (2025)'s research based on data of prefecture-level cities in China showed that the level of domestic market integration, by reducing regional trade costs, can amplify the promotion effect of trade openness on economic growth, and this effect is more obvious in the southern region.

2.2 The Role of Control Variables in the Trade Openness - Economic Growth Relationship

The selection and handling of control variables are crucial factors affecting the reliability of empirical results. Existing literature generally focuses on the moderating effects of variables such as financial development, government size, and human capital.

In terms of financial development, the mainstream view holds that a well-developed financial system can provide financing support for trade activities and reduce transaction costs, thereby strengthening the growth effect of trade openness (Tarchoun & Ghraieb, 2022). However, some studies have also found that if there are resource mismatches or excessive virtualization in financial development, it may lead to funds leaving the real economy, thereby inhibiting economic growth (Nguyen et al., 2026). Regarding the impact of government size, there is controversy: moderate government spending can provide public goods and improve infrastructure, creating a favorable environment for trade openness; but excessive government intervention may lead to resource misallocation and efficiency decline, weakening the growth effect of trade openness (Cenaj & Kilaj, 2025). In terms of human capital, Lee & Zhao (2025) pointed out that highly skilled labor can better absorb the advanced technologies and management experience brought by trade openness, thereby promoting technological progress and economic growth. However, insufficient or improper education quality may prevent the full release of the human capital effect.

2.3 Evolution and Controversies of Research Methods

The choice of empirical research methods directly affects the credibility of the conclusions. Early studies mostly employed

the mixed OLS model, but due to the failure to control for country heterogeneity and time shocks, the estimated results might be biased (Morán et al., 2025). With the development of panel data econometric methods, the fixed effect model has been widely applied. By controlling for country characteristics that do not change over time and time factors that do not vary across countries, it can more accurately identify the net effect of the core explanatory variables (Yu & Gu, 2025). Additionally, the methods of robustness tests have become increasingly diverse, including core variable substitution, extreme value handling, and instrumental variable methods, which are used to verify the reliability of the conclusions (Ayesu et al., 2026).

However, there are still some gaps in the existing research: Firstly, some studies did not fully compare the impact of including or excluding control variables on the results, possibly ignoring the problem of omitted variable bias; Secondly, the analysis of the influence of the choice of fixed effect type is insufficient, failing to clearly distinguish the independent effects of country heterogeneity and time shocks; Thirdly, some studies have a single robustness test method, and the reliability of the conclusions needs further verification. Based on this, this paper conducts a systematic design of "comparison experiments with or without control variables" and "comparison of fixed effect types", and adopts dual robustness tests, aiming to make up for the shortcomings of existing research and provide more reliable empirical evidence.

3. Research Design

3.1 Model Specification

To gradually examine the impact of trade openness on economic growth and to compare the effects of control variables and fixed effects, this paper constructs the following three types of empirical models:

3.1.1 Baseline OLS Model

To examine the impact of including control variables on the estimation results, two sets of baseline OLS models were designed:

Model 1 (without control variables): Only includes the core explanatory variables and the constant term, to test the pure correlation between trade openness and economic growth

$$gdp_g_{it} = \alpha_0 + \alpha_1 trade_{it} + \varepsilon_{it}$$

Model 2 (with control variables): Incorporate all control variables to identify the net effect of trade openness

$$gdp_g_{it} = \alpha_0 + \alpha_1 trade_{it} + \alpha_2 credit_{it} + \alpha_3 gov_{it} + \alpha_4 edu_{it} + \varepsilon_{it}$$

3.1.2 National Fixed Effects Model

To control for the heterogeneity of countries that does not change over time (such as institutional environment, resource endowment, etc.), the model is set up as follows:

$$gdp_g_{it} = \beta_0 + \beta_1 trade_{it} + \beta_2 credit_{it} + \beta_3 gov_{it} + \beta_4 edu_{it} + \mu_i + v_{it}$$

Among them, μ_i represents the fixed effect of the country, and v_{it} represents the random error term.

3.1.3 Two-way Fixed Effects Model

While controlling for country heterogeneity and time shocks (such as the global economic cycle, international policy changes, etc.), the model is set up as follows:

$$gdp_g_{it} = \beta_0 + \beta_1 trade_{it} + \beta_2 credit_{it} + \beta_3 gov_{it} + \beta_4 edu_{it} + \mu_i + \lambda_t + \omega_{it}$$

Among them, λ_t represents the fixed effect of the year, and ω_{it} represents the random error term.

3.2 Variable Definition and Measurement

Table 1 Variable Definition and Measurement

variable	variable symbol	measurement mode	Data sources
GDP growth rate	gdp_g	Annual GDP growth rate (compared to the previous year) (%)	World Bank's WDI database
trade openness	trade	The proportion of total import and export volume to GDP (%)	World Bank's WDI database
Export openness	export	The proportion of total exports to GDP (%)	World Bank's WDI database

variable	variable symbol	measurement mode	Data sources
financial development	credit	Domestic credit to private sector (% of GDP)	World Bank's WDI database
government scale	gov	General government final consumption expenditure (% of GDP)	World Bank's WDI database
human capital	edu	School enrollment, tertiary (% gross)	World Bank's WDI database

3.3 Data Sources and Processing

The panel data used in this study covers the period from 2010 to 2023, spanning 13 years. The cross-sectional individuals consist of 97 countries worldwide (including high-income, middle-income, and low-income countries, ensuring sample representativeness), with a total of 1,100 observations. The data mainly come from the World Bank's WDI database (World Development Indicators), which is widely adopted in cross-national empirical research (Cenaj & Kilaj, 2025; Nguyen et al., 2026).

During the data processing, the following main tasks were carried out: (1) Data cleaning: Remove observations with missing values for the core variables (gdp_g, trade, credit, gov, edu) to ensure the completeness of the regression sample; (2) Extreme value handling: In the robustness test, perform 1% two-way truncation on the dependent variable and the core explanatory variables, replace extreme outliers, and avoid their interference with the regression results; (3) Descriptive statistics: Test the distribution characteristics of the variables, and the results show that there are no serious abnormalities in each variable, which can be used for subsequent regression analysis.

4. Analysis of Empirical Results

4.1 Descriptive Statistics

Table 2 presents the descriptive statistics of all variables. From the distribution of the data, the mean of the explained variable gdp_g is 3.0686%, the standard deviation is 3.8162%, the minimum value is -17.8212%, and the maximum value is 24.624%, indicating that the economic growth of various countries in the sample varies significantly, including both emerging economies with rapid growth and countries facing recession; the mean of the core explanatory variable trade is 96.6172%, the standard deviation is 64.4280%, and the maximum value reaches 442.62%, while the minimum value is only 22.4862%, reflecting the significant differences in the degree of trade openness among various countries, which conforms to the actual characteristics of the global trade pattern; in terms of control variables, the means of credit, gov, and edu are 70.6070%, 16.9719%, and 54.3505% respectively, and the standard deviations are relatively large, indicating that there are obvious heterogeneities among different countries in terms of financial development, government intervention, and human capital levels, providing a data basis for testing the influence of control variables.

Table 2 Descriptive Statistics of Variables

variable	count	mean	std	min	Max
gdp_g	1100	3.0686	3.8162	-17.8212	24.6240
trade	1100	96.6172	64.4280	22.4862	442.6200
export	1100	47.3276	35.0398	5.1192	222.3085
credit	1100	70.6070	46.2217	9.3277	264.4421
gov	1100	16.9719	5.0477	5.0393	41.8557
edu	1100	54.3505	26.8980	2.1391	166.3896

4.2 Baseline OLS Regression: Comparison of Presence or Absence of Control Variables

Table 3 presents the comparative results of the baseline OLS regression, focusing on the impact of including or not including control variables on the coefficient of trade openness. Column (1) shows the regression result without including control variables, indicating that the coefficient of trade is 0.0015, with a standard error of 0.002. It fails to pass the significance

test at the 10% level, and the model R^2 is only 0.001, suggesting that when only considering trade openness, its explanatory power for economic growth is extremely weak. This might be due to the omission of key influencing factors such as financial development and government size, resulting in the concealment of the core effect.

Column (2) shows the regression result after including all control variables. It indicates that the coefficient of trade rises to 0.0054, with a standard error of 0.002, being significantly positive at the 1% level. The model R^2 also increases from 0.001 to 0.1220, and the F statistic is significant at the 1% level, indicating that the inclusion of control variables not only makes the growth effect of trade openness more apparent but also significantly enhances the overall explanatory power of the model. From the perspective of control variables, the coefficient of credit is -0.0109 (standard error = 0.003, $p < 0.01$), the coefficient of gov is -0.1677 (standard error = 0.024, $p < 0.01$), and the coefficient of edu is -0.0153 (standard error = 0.005, $p < 0.01$), all being significantly negative at the 1% level. This initially suggests that some countries within the sample period may have problems such as financial resource misallocation, excessive government intervention, and insufficient education quality.

Table 3 Baseline OLS Regression Results

variable	(1) Only includes the core explanatory variables	(2) Include all control variables
trade	0.0015	0.0054***
	(0.0020)	(0.0020)
credit	—	-0.0109***
		(0.0030)
gov	—	-0.1677***
		(0.0240)
edu	—	-0.0153***
		(0.0050)
const	2.9417***	7.0000***
	(0.1930)	(0.4110)
R^2	0.0010	0.1220
F-statistic	0.775	37.860***
No. Observations	1352	1100

Notes: *** , ** and * denote significance at 1%, 5%, and 10% levels, respectively.

4.3 Fixed Effect Regression: National FE vs Two-way FE

Table 4 compares the regression results of controlling only for national fixed effects and simultaneously controlling for both national and year-specific fixed effects. Both models include all control variables and focus on analyzing the impact of heterogeneity treatment on the estimation results. Column (1) shows the result of controlling only for national fixed effects, where the coefficient of trade significantly increases to 0.0642, with a standard error of 0.0094, being statistically significant at the 1% level, R^2 is 0.1751, and the F-statistic is also significant at the 1% level, indicating that after controlling for the heterogeneity of countries that do not change over time, the positive effect of trade openness on economic growth is fully released, and the model's explanatory power for the time fluctuations in economic growth is also further enhanced.

Column (2) shows the regression result of two-way fixed effects, where the coefficient of trade is 0.0386, with a standard error of 0.0076, still being statistically significantly positive at the 1% level, but the absolute value of the coefficient is lower than that of the national fixed effects model. R^2 is 0.1531, and the F-statistic is also significant at the 1% level. This result indicates that time shocks (such as the aftermath of the 2008 financial crisis, the 2020 COVID-19 pandemic, etc.) will slightly suppress the growth effect of trade openness, but it does not change its positive core conclusion, and the two-way fixed effect model, which simultaneously controls for country heterogeneity and time shocks, has more reliable estimation results.

From the control variables, credit is significantly negative in both models, and the absolute value of the coefficient is close (-0.0587 vs -0.0603), with standard errors of 0.0091 and 0.0069 respectively, indicating that the negative impact of financial development is stable; the coefficient of gov is -0.4027 (standard error = 0.0709, p<0.01) in the national fixed effects model and drops to -0.1604 (standard error = 0.0556, p<0.01) in the two-way fixed effects model, still being significantly negative but with a weakened intensity; the coefficient of edu is -0.0120 (standard error = 0.0146, p>0.1) in the national fixed effects model and turns positive (0.0073) in the two-way fixed effects model but is not statistically significant (standard error = 0.0133), indicating that the impact of human capital has model dependence.

Table 4 Regression Results of Fixed Effects

variable	(1) Only control for the fixed effects of the country	(2) Two-way fixed effect
trade	0.0642***	0.0386***
	(0.0094)	(0.0076)
credit	-0.0587***	-0.0603***
	(0.0091)	(0.0069)
gov	-0.4027***	-0.1604***
	(0.0709)	(0.0556)
edu	-0.0120	0.0073
	(0.0146)	(0.0133)
const	—	—
R^2	0.1751	0.1507
F-statistic	53.009***	43.734***
National fixed effect	controlled	controlled
Year fixed effect	Uncontrolled	controlled
No. Observations	1100	1100

Notes: *** , ** and * denote significance at 1%, 5%, and 10% levels, respectively.

4.4 Robustness Test

To verify the reliability of the core conclusion, this paper employs two robustness testing methods. The results are shown in Table 5. Both types of tests are based on the two-way fixed effects model framework, ensuring consistency with the baseline model setup.

Column (1) is the core variable robustness test. The export is used to replace the original core explanatory variable, trade openness (trade). The result shows that the coefficient of export is 0.0593, with a standard error of 0.0147, which is significantly positive at the 1% level and consistent with the positive effect of the original core variable; among the control variables, credit is significantly negative (-0.0606, standard error = 0.0070), gov is significantly negative (-0.1242, standard error = 0.0578), and edu is positive but not significant (standard error = 0.0134), which is consistent with the result characteristics of the two-way fixed effect baseline model, indicating that the core conclusion is not affected by the measurement method of the explanatory variables.

Column (2) is the 1% winsorization test. The dependent variable (gdp_g) and the core explanatory variable (trade) are winsorized at the 1% level on each tail, resulting in the processed variables gdp_g_wins and trade_wins. The regression results show that the coefficient of trade_wins is 0.0407, with a standard error of 0.0075, which is significantly positive at the 1% level and close to the coefficient of the original core variable in the two-way fixed effect model (0.0386); the coefficient symbols and significance of the control variables have not fundamentally changed, indicating that the core conclusion is not affected by extreme values.

Table 5 Results of Robustness Test

variable	(1) Core variable replacement	(2) Winsorized at the 1% level
export	0.0593***	—
	(0.0147)	
trade_wins	—	0.0407***
		(0.0075)
credit	-0.0606***	-0.0560***
	(0.0070)	(0.0064)
gov	-0.1242**	-0.1318**
	(0.0578)	(0.0512)
edu	0.0095	0.0007
	(0.0134)	(0.0123)
R^2	0.1429	0.1535
National fixed effect	controlled	controlled
Year fixed effect	controlled	controlled
No. Observations	1100	1100

Notes: *** , ** and * denote significance at 1%, 5%, and 10% levels, respectively.

5. Interpretation of Empirical Results

5.1 The Key Role of Including Control Variables: From "Effect Concealment" to "Real Manifestation"

The comparative results of the baseline OLS regression clearly show that the inclusion of control variables is crucial for identifying the true growth effect of trade openness. When control variables are not included, the trade coefficient is not significant and the model explanatory power is extremely weak. This might be because variables such as financial development and government size have interactions with trade openness. If not controlled, it will lead to omitted variable bias, thereby masking the true effect of trade openness. After including control variables, the trade coefficient is significantly positive and the model explanatory power improves, indicating that the positive impact of trade openness on economic growth is an independent net effect independent of financial development, government size, and human capital. This result is also consistent with the research conclusion of Cenaj & Kilaj (2025), that is, after fully controlling for interfering factors, the growth effect of trade openness can be accurately identified.

5.2 The Impact of Fixed Effect Types: The "Amplification Effect" of Country Heterogeneity and the "Slight Inhibition" of Time Shocks

The comparative results of the fixed effect regression reveal the differentiated impacts of country heterogeneity and time shocks on the growth effect of trade openness. When only controlling for country fixed effects, the trade coefficient is significantly higher than that of the baseline OLS and two-way fixed effect models, indicating that country heterogeneity (such as institutional environment, resource endowment, infrastructure level, etc.) is an important prerequisite for releasing the growth dividend of trade openness. In countries with improved institutions and good infrastructure, trade openness can more smoothly promote growth through channels such as technology spillovers and deepening of division of labor; while in countries with lagging institutions and weak infrastructure, the effect of trade openness may be counteracted (Ayesu et al., 2026). After controlling for time shocks, the trade coefficient slightly decreases but remains significantly positive, indicating that common short-term time factors such as the global economic cycle and the pandemic will have a short-term suppression on the growth effect of trade openness, but will not change its long-term positive trend. This finding also provides a perspective for understanding the fluctuations in the growth effect of trade openness during the period of anti-globalization.

5.3 Analysis of the Effects of Control Variables: The Negative Impact of Financial Development and Government Size

The regression results of the control variables show that financial development (credit) is significantly negative in most models, which is inconsistent with the expectations of traditional financial development theory. The possible reason is that during the sample period, some countries (especially developing countries) had problems of financial resource misallocation - the financial system tended to allocate credit resources to speculative fields such as real estate and stock markets rather than small and medium-sized enterprises and innovative enterprises in the real economy, resulting in financial development failing to effectively support trade openness and economic growth, and instead exacerbating economic fluctuations (Tarchoun & Ghraieb, 2022). The negative impact of government size (gov) indicates that in some countries in the sample, the government intervention was excessive, public expenditure efficiency was low, and there were problems such as corruption and rent-seeking, leading to resource misallocation and suppressing market vitality and the release of trade openness dividends (Cenaj & Kilaj, 2025).

The effect of human capital (edu) is model-dependent. This result is consistent with the research of Lee & Zhao (2025), that is, the impact of human capital on economic growth not only depends on the number of education , but also on the quality and allocation efficiency of education. Although some countries in the sample have a high school enrollment rate, the educational content is disconnected from industrial demands, and highly educated workers engage in low-skilled jobs, resulting in a low marginal output of human capital and an inability to fully manifest the effect.

5.4 The Significance of Robustness Tests: Verification of the Reliability of Core Conclusions

The results of the dual robustness tests further confirm the reliability of the core conclusion of this paper. The core variable substitution test shows that regardless of whether the openness is measured by the proportion of import and export or the proportion of export, its positive effect on economic growth is significant and is not affected by the measurement method of the variable; the 1% winsorization test excludes the distortion of the regression results caused by extreme values, indicating that the growth effect of trade openness is universal and not driven by individual abnormal samples.

6. Conclusion and Policy Recommendations

6.1 Conclusion

The degree of trade openness has a significant and stable positive impact on economic growth. This conclusion holds true under various model settings, including the inclusion of control variables, controlling for national heterogeneity and time shocks, replacing core variables, and handling extreme values. It confirms the long-term growth dividend of trade openness.

The inclusion of control variables and the handling of fixed effects are crucial for the estimation results: Failure to include control variables would mask the true effect of trade openness. After controlling for national heterogeneity, the effect intensity significantly increases. Time shocks slightly weaken but do not change the core conclusion.

Financial development and government size have a negative impact in most models, reflecting that some countries have problems such as financial resource misallocation and excessive government intervention. The effect of human capital has model dependence, and the release of its growth dividend depends on the improvement of education quality and allocation efficiency.

6.2 Policy Recommendations

Maintain a trade liberalization orientation and unleash the benefits of openness. All countries should abandon trade protectionism, actively reduce trade barriers, expand the scale of imports and exports, especially developing countries should actively integrate into the global trading system, absorb advanced technologies and management experience through trade openness, and improve the efficiency of resource allocation. At the same time, attention should be paid to optimizing the trade structure, avoiding excessive reliance on low value-added product exports, and enhancing the sustainability of trade openness.

Optimize the financial system structure and enhance the ability to serve the real economy. In response to the negative impacts of financial development, all countries should strengthen financial supervision, guide credit resources to the real economy, especially small and medium-sized enterprises and innovative enterprises; improve the financial market system, enhance the accessibility and efficiency of financial services, develop inclusive finance, and avoid excessive concentration of financial

resources in speculative fields; developing countries should gradually promote financial marketization reforms, reduce financial fragility.

Reasonably control the size of the government and enhance public governance efficiency. All countries should optimize the structure of government expenditures, reduce unnecessary administrative intervention, and focus on increasing investment in infrastructure, education, and healthcare public products; enhance government governance capabilities, strengthen anti-corruption and integrity building, and improve the efficiency of public expenditure; through streamlining administration and delegating powers, optimizing the business environment, fully stimulate market vitality, and create conditions for the release of trade openness benefits.

Strengthen the quality of human capital and promote effective transformation. All countries should increase investment in education, not only focus on improving the coverage rate of basic education, but also pay attention to the matching of education quality and industrial demands; strengthen the connection between vocational education and higher education, cultivate skilled talents that meet market demands; improve the mechanism for allocating human capital, reduce talent mismatches, and fully utilize the ability of human capital in absorbing the spillover effects of trade openness.

Take into account the heterogeneity characteristics of one's own country and formulate differentiated policies. Developing countries should focus on improving the institutional environment and infrastructure to lay a foundation for the release of trade openness benefits; developed countries should pay attention to the synergy between trade openness and innovation, high-end manufacturing, and further strengthen the spillover effect of technology; all countries need to pay attention to the impact of global time shocks, flexibly adjust trade policies and macro-control measures, and cope with external environmental fluctuations.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Ayesu, E. K., Sakyi, D., & Arthur, E. (2026). Trade openness and economic growth in Africa: The role of seaport efficiency. WMU Journal of Maritime Affairs, 1–20.

[2] Cenaj, A., & Kilaj, D. (2025). The impact of trade openness on the economic growth of Southeast European countries. SEEU Review, 20(2), 3–19.

[3] Lee, C. C., & Zhao, Y. N. (2025). What affects national innovative capacity? The role of economic growth and trade openness. Economic Analysis and Policy, 86, 1099–1118.

[4] Morán, R. O., Llonto, Y., Cúneo, F. E., Lizana, N. P., Castro, P. J., Pérez, M. J., & Montenegro, M. E. (2025). The macroeconomic determinants of trade openness in Latin American countries: A panel data analysis. F1000Research, 13.

[5] Nguyen, T. V., Dinh, T. K. C., Vu, N. X., & Pham, X. H. (2026). Determinants of economic growth in Korea: The roles of financial development, energy use, trade openness, and open innovation (ARDL evidence). Energy Conversion and Management: X, 30, 101823.

[6] Pomerlyan, E., Belitski, M., Orel, A., & Bondarenko, A. (2026). Trade openness in economic performance: a comparative perspective of the world integration blocs. Eurasian Economic Review, 1–51.

[7] Tarchoun, M., & Ghraieb, I. (2022). Trade openness relationship with macroeconomic factors on Saudi Arabian economy: ARDL application. International Trade, Politics and Development, 6(3), 121–129.

[8] Trabelsi, R., Kalai, M., & Helali, K. (2025). The nonlinear impact of trade openness on the economic growth in Maghreb countries. Quality & Quantity, 60(1), 1–22.

[9] Yu, C., & Gu, C. (2025). Trade openness, market integration and economic growth: Evidence from China. International Review of Financial Analysis, 106, 104484.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

CEO Compensation Contract Design under Moral Hazard and Adverse Selection

Chenlu Fang*

School of Finance, Shanghai University of International Business and Economics, Shanghai, 201620, China

Corresponding author: Chenlu Fang

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: In modern corporate governance, the separation of ownership and management rights leads to increasingly prominent principal-agent problems. Information asymmetry leads to adverse selection in advance and moral hazard after the event, which significantly increases agency costs and weakens the incentive effect of enterprises. In view of the realistic dilemma of unobservable executive ability and unverifiable efforts, the traditional single performance-linked compensation model is difficult to adapt to the differentiated needs of different types of executives. Therefore, this study focuses on the design of CEO compensation contract, and constructs a theoretical model that includes both moral hazard and adverse selection. It aims to achieve type separation and incentive compatibility through mechanism design, so as to maximize shareholders ' interests under the condition of information asymmetry. Based on the principal-agent theory framework, this study sets shareholders as risk-neutral and CEOs as risk-averse agents, and constructs a single-period model including hidden ability types and hidden effort levels. By introducing a linear compensation contract and a mean square utility function, the CEO 's optimal effort decision is internalized, and a contract menu is constructed based on participation constraints and incentive compatibility constraints. Aiming at the problem that the model is difficult to solve analytically, a two-stage grid search method is used for numerical optimization. Firstly, the high-yield area is located by global scanning of coarse grid, and then the optimal incentive coefficient and fixed compensation combination are determined by local fine search of fine grid. This method effectively improves the efficiency and stability of the solution, and has good expansibility. The results show that the optimal compensation contract shows significant differentiation characteristics : high-capacity CEOs correspond to high performance sensitivity coefficient (b _ H * = 0.837) and negative fixed compensation (w _ H * = -0.638), forming a ' strong incentive and high deposit ' model ; the CEO with low ability corresponds to a lower incentive intensity (b _ L * = 0.327) and a fixed salary close to zero. This design effectively compresses the information rent of high-ability CEOs, and at the same time promotes the two types of CEOs to truthfully reveal their own types and make corresponding efforts to achieve incentive compatibility and type self-selection. This study not only expands the contract theory under dual information asymmetry, but also provides theoretical basis and practical reference for enterprises to design differentiated executive compensation and regulators to formulate reasonable incentive policies. In the future, it can be further extended to continuous type setting, multi-task performance indicators and dynamic contract scenarios to enhance the realistic explanatory power of the model.

Keywords: Moral Hazard; Adverse Selection; Salary Contract Design; Principal-Agent Model

Published: May 20, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1459

1. Introduction

1.1 Research Background

The design of executive compensation contract is one of the core issues of corporate governance. In modern companies with the separation of ownership and management rights, due to the information asymmetry caused by principal-agent, the ex ante adverse selection (the ability information of executives is opaque and cannot choose excellent talents) and the ex post moral hazard (the behavior of executives cannot be observed, which maximizes their private interests and harms the interests of shareholders), resulting in an increase in the agency cost of the company and a reduction in the incentive effect of the company 's behavior.

In theory, contract research has gradually developed from unilateral information asymmetry to bilateral information asymmetry. Early research mainly focused on the moral hazard and adverse selection problems in the case of unilateral information asymmetry, and proposed solutions such as risk sharing mechanisms and menu contracts. Subsequently, Picard et al. concerned about the contract design of two risk concurrent situations. In the case of two kinds of information asymmetry, the contract structure is more stringent and more challenging. In recent years, Chinese scholars have also begun to study the dual agency problem, and gradually try to design the corresponding incentive model for the two agency problems.

In practice, facing the dilemma of unobservable ability and unverifiable efforts, it has always been a problem to adopt which salary system or plan to deal with. Generally, the performance-linked approach (equity, bonus) is adopted, but for different levels of executives, there is no set of incentives to adapt to all executives. To this end, we must design a differentiated compensation contract. On the one hand, it is necessary to induce executives to choose their own programs, on the other hand, it is necessary to encourage executives to work as hard as possible to complete the task, which requires the design of compensation contracts to maximize the interests of shareholders and executives on the basis of ensuring this premise. Incentive compatibility.

Figure 1: Research background



1.2 Research Objectives and Significance

In view of the above background, the goal of this study is to explore the optimal design method of CEO compensation incentive contract under the framework of moral hazard and adverse selection. Specifically, this study attempts to answer: when executives have different levels of ability and their ability types are not known to the principal in advance, and the degree of effort is not observable afterwards, how should shareholders set fixed compensation and performance sensitivity coefficients to simultaneously motivate different types of executives to work hard and make them reveal their own types truthfully, so as to maximize corporate performance ?

In order to achieve this goal, this study establishes a theoretical model containing ' double asymmetric information ', introduces risk-averse managers and risk-neutral shareholders, and comprehensively uses principal-agent theory and numerical optimization methods to solve the optimal compensation contract and analyze its economic meaning.

The research has the following significance:

First, it enriches the theoretical analysis paradigm of executive compensation incentives. In the past, the research on executive compensation contracts mostly focused on a single information asymmetry problem or empirical correlation analysis. This study applies the mechanism design theory to executive incentive contracts, and considers both hidden information and hidden actions in the model, which has reference value for the expansion of classical contract theory in the context of corporate governance.

Second, the model solving method of this study is general and can be extended to other dual principal-agent problems. For example, in the face of multi-type employees and hidden efforts, the internal incentive design of enterprises, the phased incentive contract design of venture capitalists and entrepreneurs, etc., only need to adjust the model parameters to apply this research method to solve, which is helpful to improve the numerical stability and efficiency of solving such problems.

Thirdly, the conclusions of this study have guiding significance for practice. Through model simulation, this study clarifies the composition characteristics of optimal contracts for executives with different abilities, and provides a quantitative basis for enterprises to design differentiated compensation incentives based on managers ' abilities. At the same time, the analysis of this study reveals the trade-off relationship between information rent and effort incentive in executive incentive, which can provide reference for regulators to formulate executive compensation policies and improve corporate governance.

1.3 Contributions of This Study

The main contributions of this study include :

Firstly, an executive incentive contract model with both moral hazard and adverse selection is built. Considering the heterogeneity of executive ability, the influence mechanism of the optimal contract under double information asymmetry is described. Compared with the single information asymmetry model, it is more in line with the real corporate environment and makes up for the lack of domestic analysis of the dual principal-agent theory of executive compensation.

Secondly, using the novelty of the solution method, this study combines analytical derivation and numerical search. Based on the analysis of the set of bindings under constraints, the original complex constrained optimization problem is transformed into a search problem in a limited parameter space. Then, the optimal solution is obtained by using the method of coarse grid local search and fine grid local refinement, which improves the accuracy and efficiency of the solution, and provides a new technical means for the design of similar incentive contracts.

Thirdly, combined with the results of the model, it can be seen that the optimal contract is a differentiated model of ' strong incentive and high ability, weak incentive and low ability ' for different types of ability managers. Under this new model, although high ability managers obtain high performance coefficient, they get more effective information rent compression. The incentive intensity of low-ability managers is reduced, so that high-ability and low-ability managers can maintain the true type, which also shows that the incentive marginal income for low-ability managers is more stable. This part uses the method of model to expand the classical principal-agent theory, changes the traditional incentive method to a certain extent, and explains the problem of executive incentive in reality.

To sum up, the exploration of the theory and method of this study provides a certain reference significance for improving the research of executive compensation contract and guiding enterprise practice.

2. Related Work

Due to the widespread existence of principal-agent relationship, the academic circles have carried out a lot of research on incentive contracts under the condition of information asymmetry. According to the different types of information asymmetry, the relevant literature can be divided into the following aspects :

(1) Research on incentive contract with only moral hazard. This kind of research assumes that the agent 's private information is only an after-action, that is, the effort is unobservable, but its ability type, task difficulty, etc. are known in advance. For example, Holmstrom (1979) established a single-period hidden action model to prove that the optimal contract needs to balance risk sharing and incentive provision, usually manifested as allowing the agent to bear part of the output risk[1]. Shavell (1979) further analyzed the form of optimal incentive compensation, such as linear contract can achieve suboptimal incentive effect under certain conditions. Domestically, Wang Xianyu and Yan Jinjiang discussed the combination of incentives and supervision, and Hong Jianqiao compared the incentive effects of the compensation scheme based on budget indicators and

the linear compensation scheme. These studies focus on how to promote the agent 's diligent work through contract design in the case of unobservable efforts, forming a perfect theoretical system of ' type I agency problem '.

(2) Only discuss the incentive contract of adverse selection. This kind of literature analyzes how the principal designs a mechanism to encourage the agent to voluntarily disclose the true type of information when some of the agent 's private information (such as ability, high-quality opportunities) is not known to the principal before the contract is signed. Spence (1973) pioneered the study of adverse selection by giving the classical signal transmission model. Based on the mechanism design framework, Myerson (1982) theoretically gave the optimal auction and contract under adverse selection. In the principal-agent environment, Mirrlees (1976), Maskin and Riley (1984) and others studied the optimal incentive compatibility contracts of single continuous type and multiple continuous types respectively. It is found that when there is a high type of high information rent, it is in line with the optimal contract requirements to sacrifice some low types in order to expose the identity of the high type, which can also be regarded as ' the information rent obtained by high-ability executives '. This corresponds to the incentive mechanism of executives in the optimal contract mentioned above. It is similar to the research conclusion of domestic scholar Zhang Xinli[2] and others on venture capital contract, that is, the optimal separation contract can induce entrepreneurs with different abilities to self-select ; in general, the ' menu ' contract design is only applicable to a single implicit information situation.

(3) Research on incentive contract with moral hazard and adverse selection. When the agent has both pre-hidden information and post-hidden action, the incentive contract design needs to meet two types of incentive compatibility conditions at the same time[3], and the problem becomes more complicated. There are relatively few studies on this situation and most of them analyze special cases in theory. One idea is to find sufficient conditions to ensure that both types of constraints are satisfied. For example, the first-order rule proposed by Rogerson (1985) can be applied to the coexistence of adverse selection and moral hazard under certain conditions. Picard (1987) is one of the important works in this field. He analyzed the general properties of the optimal incentive mechanism in the presence of both hidden information and hidden actions. Later, Theilen (2003) and Sung (2005) extended the model situation, including the dual information asymmetry problem under continuous time and dynamic settings. Ding Yuanyao[4], a domestic scholar, proposed the " dual incentive and restraint mechanism " earlier, while Zhang Wei[5] focused on the linearization and simplification of the contract form, so as to obtain an enforceable linear incentive contract when the ability difference induces double information asymmetry. Niu Xiaoqin et al. further considered the influence of the agent 's social justice preference factor on the incentive contract. Generally speaking, these literatures provide ideas for understanding contracts under dual information asymmetry, but they often assume that the model is simple to obtain analytical solutions, or use numerical examples to verify specific conclusions. There is still a lack of systematic solution to the optimal contract under general parameters [6]. This study has made efforts in this regard. This study will use numerical methods to solve the optimal contract in general situations, and apply the theory to the specific scenario of CEO compensation design, which will form a useful supplement to related research.

3. Theoretical Model and Setting

3.1 Economic Environment and Information Structure

This study considers a single-period principal-agent model including shareholders (principals) and CEOs (agents). Shareholders own the ownership of the company and care about the company 's profits, and the CEO is responsible for business decisions and efforts. The information structure of the two sides is asymmetric : CEO has two types of capabilities, denoted by $\theta \in \theta_L, \theta_H$, which represent ' low ability ' and ' high ability ' types, respectively. In the case of high ability, $\theta_H > \theta_L > 0$. θ determines the efficiency or marginal output capacity of CEO 's efforts to transform into performance. The adverse selection is that θ is only known to the CEO himself before the contract is signed, and shareholders can only know a prior probability $P(\theta=\theta_H)=p$ (the proportion of high-capacity managers), but they cannot directly identify the true type of a CEO candidate. In addition, the CEO chooses the level of effort $e \geq 0$ to operate the company during the execution of the contract, and the effort cannot be observed by shareholders, reflecting moral hazard. CEO 's effort will affect the company 's final operating performance or profit, expressed as a random variable. In order to simplify the analysis, this study assumes that the company 's final profit Π satisfies the following production function:

$$\Pi=\theta \cdot e+\varepsilon \tag{3-1}$$

where ε represents independent random noise, reflecting the uncertainty of the operating environment. Let $E[\varepsilon]=0, Var(\varepsilon)=\sigma^2$. This hypothesis implies that the contribution of CEO effort to expected profit is θe, and the equal effort of high-capacity CEO will bring higher expected performance. At the same time, due to the noise, the actual performance contains uncontrollable components. In order to highlight the core of the problem, this study assumes that shareholders are risk-neutral, while CEOs are risk-averse and satisfy the mean square deviation preference structure. Specifically, let the CEO 's utility function be $U(w)$, where w is the actual monetary income. This study assumes that $U(w)$ is an augmented concave function, and the risk cost can be represented by mean-variance equivalence (such as exponential utility brings linear variance penalty). This setting is often used to analyze the optimization problem of linear contracts under risk aversion conditions.

Based on the above settings, there are two types of information asymmetry : pre-hidden information about θ (adverse selection) and post-hidden action about e (moral hazard). The transaction sequence is as follows : First, the shareholders propose a compensation contract menu for the CEO to choose, and each contract corresponds to a set of payment options ; the CEO chooses to accept one of the contracts or rejects according to his own ability type (if rejected, the income of both parties is an exogenous reserve value) ; after the contract is signed, the CEO chooses the effort level e according to the terms of the contract ; subsequently, the profit Π is realized and allocated to both parties according to the contract. The goal of this study is to solve the optimal compensation contract design that maximizes shareholder returns in such an environment.

3.2 Technology and Preference

Output and Cost : CEO 's effort e will bring personal costs. This study assumes that the cost function is

$$C(e)=\frac{1}{2}ke^2 \tag{3-2}$$

Where $k>0$ denotes the effort cost coefficient, and the larger k means the higher cost per unit effort. The production (profit) function is $\Pi=\theta e+\varepsilon$ as described above. A high ability θ_H means that the same effort can produce a higher expected output. The noise ε obeys a distribution with a mean value of 0 and a variance of σ^2 (e.g.normal distribution), independent of θ and e.

Contract form: This study considers linear compensation contracts, which have a certain realistic correspondence in executive compensation (such as basic annual salary plus performance bonus). Let the contract be (w,b), where w is the fixed salary, and b is the bonus coefficient (or incentive intensity) linked to performance. When the actual profit is Π, the CEO gets $w+b\Pi$, and the shareholders get $\Pi-(w+b\Pi)=(1-b)\Pi-w$. Since $0\le e\le e_{max}$ and the output is bounded, this study limits $0\le b\le 1$ to avoid over-reliance on performance ratio in contract payment. It is worth noting that if w is negative, it means that the manager invests or guarantees, which is equivalent to the performance deposit mechanism ; if $1-b<0$, the shareholders need to pay extra, this extreme case is not considered here.

Agent utility: CEO is risk-averse and faces random income. Certainty Equivalent (CE) can be expressed as expected income minus risk cost. In the form of mean square error, when CEO chooses effort e under contract (w,b), its utility is determined as follows:

$$U^{CE}=w+b \cdot E[\Pi]-C(e)-\frac{1}{2}\gamma b^2 Var(\Pi) \tag{3-3}$$

Where $\gamma>0$ is the CEO 's absolute risk aversion coefficient. Since $E[\Pi]=\theta e$, $Var(\Pi)=\sigma^2$, the output expectation and variance are substituted, and the above equation can be written as:

$$U^{CE}(e;w,b,\theta)=w+b\,\theta e-\frac{1}{2}ke^2-\frac{1}{2}\gamma b^2\sigma^2 \tag{3-4}$$

The formula shows that the expected utility of CEO is affected by fixed income w, performance coefficient b, own ability θ, effort decision e and risk cost. CEO will maximize U^{CE} when choosing efforts; shareholders will consider the CEO 's optimal response and willingness to participate under different contracts when providing contracts.

Agent 's effort decision : Under the given contract parameters (w,b) and its own type θ, the CEO chooses effort e to maximize $U^{CE}(e)$. The optimal effort response function can be obtained by solving the first derivative of the above formula with respect to e and making it zero:

$$\frac{\partial U^{CE}}{\partial e}: b\,\theta-ke=0 \implies e^*(b,\theta)=\frac{b\,\theta}{k} \tag{3-5}$$

This means that the CEO 's optimal effort level increases proportionally with the incentive intensity b and the ability θ, and decreases with the increase of the cost coefficient k. When $b=0$ (pure fixed wage contract), $e=0$, that is, the lack of incentive will make the CEO choose zero effort ; when θ is higher and b is larger, the marginal revenue $b\theta$ of CEO 's per unit effort is higher, so he is willing to invest more effort. It should be noted that the result is the connotation solution obtained under the assumption of quadratic cost and utility mean square error. If e exceeds a certain physical upper limit, e can be constrained in the model, but this problem does not occur under the benchmark parameters of this study.

From the above analysis, it can be seen that shareholders can foresee the CEO 's effort response $e^*(b,\theta)$ when designing contracts. Therefore, in the subsequent solution, this study can substitute the CEO 's effort selection and directly replace e with the optimal effort. Specifically, for a CEO of type θ, the utility after accepting the contract is determined to be equivalent to :

$$U_\theta^{CE}=w+\frac{1}{2}\frac{b^2\theta^2}{k}-\frac{1}{2}\gamma b^2\sigma^2 \tag{3-6}$$

In this study, $e^*=\frac{b\theta}{k}$ is substituted and simplified to obtain $U_\theta^{CE}=w+\frac{b^2\theta^2}{2k}-\frac{\gamma b^2\sigma^2}{2}$. This expression clearly decomposes the source of CEO utility : w is deterministic income, $\frac{b^2\theta^2}{2k}$ is the benefit of effort (net gain after deducting the cost of effort), and $-\frac{\gamma b^2\sigma^2}{2}$ is the cost of risk. When $\gamma=0$ (CEO risk neutral), the risk cost is zero, and the CEO utility is simplified to $w+\frac{b^2\theta^2}{2k}$.

When $\gamma>0$ and b is large, the risk cost will significantly reduce the utility, which reflects that too high performance incentive is detrimental to the utility of risk-averse managers in high-risk situations.

Figure 2: Modeling code

```
# ---------- 核心函数 ----------
def CE(theta, w, b, k, sigma, gamma):  3 usages
    return w + (b**2)*(theta**2)/(2.0*k) - 0.5*gamma*(b**2)*(sigma**2)

def e_star(theta, b, k):  2 usages
    return (b*theta)/k

def wL_from_binding(bL, theta_L, k, sigma, gamma, U0):  2 usages
    return U0 - (bL**2)*(theta_L**2)/(2.0*k) + 0.5*gamma*(bL**2)*(sigma**2)

def wH_from_binding(bL, bH, theta_H, k, sigma, gamma, U0, theta_L):  1 usage
    wL = wL_from_binding(bL, theta_L, k, sigma, gamma, U0)
    return wL + ((bL**2 - bH**2)*(theta_H**2))/(2.0*k) + 0.5*gamma*((bH**2 - bL**2)*(sigma**2))

def profit_type(theta, w, b, k):  2 usages
    return ((1.0 - b)*b*(theta**2))/k - w
```

3.3 Contract Space and Constraints

Under the dual information asymmetry, shareholders can consider designing a contract menu $(,b_L),(w_H,b_H)$ for CEOs who claim to be low-capacity type (θ_L) or high-capacity type (θ_H). Here, we assume that there are only two types of executive capabilities, in order to highlight the nature of mechanism design ; the model can be extended to continuous types but the analysis will be more complex. Now, this study lists the constraints that contract design needs to meet :

Rational Constraint (IR) : The contract must ensure that all types of CEOs are willing to accept and not to reject the contract. The CEO 's retention utility U_0 (external opportunity utility) is normalized to 0. Then for CEOs of type θ_i, the assigned contract (w_i,b_i) must satisfy:

$$U^{CE}\theta_i(w_i,b_i)\geq U_0=0 \tag{3-7}$$

Since the U^{CE} expression of this study has been substituted into the optimal effort, the IR constraint is actually:

$$w_i+\frac{b_i^2\theta_i^2}{2k}-\frac{\gamma b_i^2\sigma^2}{2}\geq 0,\ \ i\in L,H \tag{3-8}$$

This means that the contract should give this type of CEO at least zero utility. Usually, in the optimal contract, the IR constraint of the low-type CEO happens to be equal, that is, the low-type participation constraint is bound. The reason is that in order to save costs, shareholders tend not to leave excess rent for low-capacity managers, only to provide just the remuneration that they are willing to participate. High-type CEOs may obtain certain excess utility (information rent) because of information advantage, and their IR constraints can be strictly satisfied by inequality.

Incentive compatibility constraint (IC) : In order to achieve type separation, it is necessary to ensure that all types of CEOs choose the ' contract designed for their own type ' in the contract menu to be optimal, that is, the honest report type is higher than the false report income. For high-capacity CEOs, it is necessary to prevent them from pretending to be low-capacity to obtain different contracts ; for CEOs with low ability, it is necessary to prevent them from falsely claiming high ability. Formally, these two types of IC constraints are:

$$U^{CE}\theta_H(w_H,b_H)\geq U^{CE}(w_L,b_L) \tag{3-9}$$

$$U^{CE}\theta_L(w_L,b_L)\geq U^{CE}(w_H,b_H) \tag{3-10}$$

Substituting the utility to determine the equivalent formula can be further written as:

$$w_H+\frac{b_H^2\theta_H^2}{2k}-\frac{\gamma b_H^2\sigma^2}{2}\geq w_L+\frac{b_L^2\theta_H^2}{2k}-\frac{\gamma b_L^2\sigma^2}{2} \tag{3-11}$$

$$w_L+\frac{b_L^2\theta_L^2}{2k}-\frac{\gamma b_L^2\sigma^2}{2}\geq w_H+\frac{b_H^2\theta_L^2}{2k}-\frac{\gamma b_H^2\sigma^2}{2} \tag{3-12}$$

The right side of the above inequality represents the utility of high-capacity CEOs when they choose a contract designed for low-capacity (the utility of subscript mismatch), and the utility of low-capacity CEOs when they choose a high-capacity contract. Incentive compatibility requires that the utility of each type under their own contract is not lower than that under the other party 's contract. Usually, in the optimal solution, the IC constraint of the high-capacity type is bound and the IC constraint of the low-capacity type is relaxed. That is to say, the optimal contract arrangement will make the high-capacity CEO have no difference in impersonating the low-capacity CEO, so that shareholders can minimize the information rent given to the high-capacity CEO ; due to capacity constraints, low-capacity CEOs may be less effective even if they are given high-capacity contracts because they fail to meet performance requirements (that is, the left side is significantly larger than the right side), so there is no special constraint on low-capacity to choose high-capacity contracts. Mathematically, a high-capacity IC binding means that the above equation takes an equal sign, which can be used to eliminate a contract parameter.

3.4 Expression of the Optimal Problem

After the model setting and constraint conditions are clear, this study can express the optimal contract design problem of shareholders. The goal of shareholders (principals) is to maximize their expected returns, that is, the company 's profits minus the remuneration paid to the CEO. Under the uncertainty of the two types of managers, their expected returns are:

$$E[\Pi\text{-(CEO compensation)}]=(1-p)[(1-b_L)\,E[\Pi|\theta_L]-w_L]+p[(1-b_H)\,E[\Pi|\theta_H]-w_H] \tag{3-13}$$

Since $E[\Pi|\theta_i]=\theta_{i,}\ \ e^*(b_i,\theta_i)=\theta_i,\frac{b_i\theta_i}{k}=\frac{b_i\theta_i^2}{k}$, we can get:

$$\text{Shareholder expected return}=(1-p)[(1-b_L)\frac{b_L\theta_L^2}{k}-w_L]+p[(1-b_H)\frac{b_H\theta_H^2}{k}-w_H] \tag{3-14}$$

Sorting out each item, the shareholder 's income is composed of profit sharing and salary reduction payment in each type of situation. In this study, they are recorded as π_L and π_H respectively to represent the benefits of low-capacity and high-capacity types to shareholders:

$$\pi_i=(1-b_i)\frac{b_i\theta_i^2}{k}-w_i,\ i=L,H \tag{3-15}$$

The problem of maximizing shareholders is:

Maximization : $(1-p)\,\pi_L+p\,\pi_H$

Variables: w_L,b_L,w_H,b_H

Constraints:

$$w_L+\frac{b_L^2\theta_L^2}{2k}-\frac{\gamma b_L^2\sigma^2}{2}\geq 0 \tag{3-16}$$

$$w_H+\frac{b_H^2\theta_H^2}{2k}-\frac{\gamma b_H^2\sigma^2}{2}\geq 0 \tag{3-17}$$

$$w_H+\frac{b_H^2\theta_H^2}{2k}-\frac{\gamma b_H^2\sigma^2}{2}=w_L+\frac{b_L^2\theta_H^2}{2k}-\frac{\gamma b_L^2\sigma^2}{2} \tag{3-18}$$

$$w_L+\frac{b_L^2\theta_L^2}{2k}-\frac{\gamma b_L^2\sigma^2}{2}\geq w_H+\frac{b_H^2\theta_L^2}{2k}-\frac{\gamma b_H^2\sigma^2}{2} \tag{3-19}$$

The above constraints correspond to IR (L), IR (H), IC (H), IC (L) in turn. Based on the above analysis, this study expects that in the optimal solution, IR (L) and IC (H) will be tightly bound (equal sign), IR (H) will be relaxed (high capacity has positive rent), and IC (L) will be relaxed (low capacity has no power to pretend to be high capacity). Based on this, the two decision variables w_L and w_H can be eliminated by using the equality conditions of IR (L) and IC (H), and then the problem is transformed into the optimization of the two incentive coefficients bL and bH. Intuitively, the contract compensation wL of low-capacity managers will be lowered by shareholders to a level that they are willing to accept, and the contract of high-capacity managers is designed to make them have no income from posing as low-capacity, so as to choose truthful reports without difference. After this treatment, shareholders only need to determine the optimal combination of (b_L,b_H). Although manual derivation is feasible, it is difficult to directly solve the explicit solution due to the complexity of the equation. In the next section, this study will use numerical methods to solve the optimal b_L^*,b_H^* and further give the corresponding w_L^*,w_H^* and other optimal quantities. It is worth mentioning that in order to ensure the global optimality of the solution, this study will combine coarse-grained global search and local fine search to scan the entire (b_L,b_H) parameter space to avoid falling into the trap of local extremum.

4. Solving Method and Modeling Route

4.1 Binding Set

According to the analysis of Section 3.4, this study gives priority to the case where the low-type participation constraint and the high-type incentive compatibility constraint are equal signs, and simplifies the solution process accordingly. The specific steps are as follows : First, the IR constraints of low-capacity CEOs are equalized, that is:

$$w_L=\frac{b_L^2\theta_L^2}{2k}-\frac{\gamma b_L^2\sigma^2}{2}-U_0 \tag{4-1}$$

Since $U_0=0$, the above formula gives the functional relationship of w_L with respect to b_L. It shows that the fixed salary paid by shareholders for low-capacitated managers just makes its utility to zero. It can be seen that w_L may be positive or negative with the increase of b_L: when b_L is small, $\frac{b_L^2\theta_L^2}{2k}$ is dominant, w_L is positive and increases; when b_L is large, the risk cost $\frac{\gamma b_L^2\sigma^2}{2}$ increases faster, and w_L may decrease or even become negative. Intuitively, this means that if low-capacity CEOs are given too high a performance coefficient b_L, their required retention fixed wage w_L will be reduced or even turned into a requirement for them to pay a deposit to compensate for the loss of utility caused by risk. Similarly, taking the equal sign of the IC constraint of the high-capacity type can be solved:

$$w_H=w_L+\frac{(b_L^2-b_H^2)\theta_H^2}{2k}-\frac{\gamma(b_H^2-b_L^2)\sigma^2}{2} \tag{4-2}$$

This formula is derived from the expression of w_H about w_L,b_L,b_H by making the utility of high-capacity CEO equal in two contracts. Substituting the previous $w_L(b_L)$, the functional relationship between w_H and b_L, b_H can be obtained. Thus, the four variables of the contract (w_L,b_L,w_H,b_H) are reduced to two bL, bH. Next, only $w_L(b_L)$ and $w_H(b_L,b_H)$ are substituted into the shareholder income function and residual constraints for optimization. At this time, it is necessary to check whether the remaining low-type IC constraints are naturally satisfied, that is, to verify that the low-capacity CEO does not have the motivation to impersonate the high-capacity contract under the optimal solution. If it is found that low-type ICs have a tendency to default, it is necessary to adjust the binding constraint set to re-solve. However, under most parameters, the constraint binding pattern assumed in this study will appear. This study will test this after solving. The whole process is actually the process of assuming the constraint binding set first, then simplifying and solving the problem by back substitution elimination, and then verifying the hypothesis. Due to the complexity of the double asymmetry problem, this study adopts

this simplified solution path based on theoretical inference, which greatly reduces the difficulty of numerical optimization.

4.2 Numerical Optimization Implementation

Because the analytical solution is extremely complex, this study mainly uses computer-aided numerical methods to solve the optimal contract. Since the objective function (shareholder expected return) is defined on a two-dimensional space, there may be multiple local extreme points in this space. Therefore, the author uses a two-stage grid search method : In the first stage, a coarse grid search is performed on (b_L,b_H) in the entire $[0, 1]\times[0, 1]$ interval to find the global high value area of the expected return ; in the second stage, on the basis of the first stage, a fine grid search is performed on the above high-value areas to obtain the optimal point. Through this method, the corresponding optimal contract can be quickly found when the coefficient value is certain.

Figure 3: Numerical optimization implementation

```
# ---------- 粗网格 ----------
b_grid = linspace_grid( a: 0.0,  b: 1.0, coarse_step)
EP_grid = np.zeros((len(b_grid), len(b_grid)))
best_coarse = None

for i, bL in enumerate(b_grid):
    for j, bH in enumerate(b_grid):
        out = objective_and_components(bL, bH, params)
        EP_grid[i, j] = out["EP"]
        key = (out["EP"], bL, bH, out)
        if (best_coarse is None) or (key[0] > best_coarse[0]):
            best_coarse = key

EP_best_coarse, bL_c, bH_c, out_c = best_coarse

# ---------- 细网格 ----------
bL_min = max(0.0, bL_c - fine_radius)
bL_max = min(1.0, bL_c + fine_radius)
bH_min = max(0.0, bH_c - fine_radius)
bH_max = min(1.0, bH_c + fine_radius)

bL_fine = linspace_grid(bL_min, bL_max, fine_step)
bH_fine = linspace_grid(bH_min, bH_max, fine_step)

EP_fine = np.zeros((len(bL_fine), len(bH_fine)))
best_fine = None
for i, bL in enumerate(bL_fine):
```

Rough grid global search: Set the step size to a larger value Δ_1 (for example, 0.01), and generate b_L, b_H as a grid point set from 0 to 1 ; enumerate all (b_L,b_H) combinations, use the formula and binding constraints in Chapter 3, Section 4 to obtain w_L and w_H, and then obtain the shareholder 's expected return $EP(b_L,b_H)$; the grid point corresponding to the maximum benefit of the record is the coarse search optimization. Due to the large step size, the whole optimal position range is mainly found at this stage.

Figure 4 Heat map of shareholders ' expected return EP with respect to (b_L,b_H) under coarse grid. The brighter the color, the greater the profit, and the blue dots mark the maximum value obtained by the rough search (rough optimal). It can be seen that EP has unimodality in the parameter space, and the color is the brightest near the optimal solution, indicating that there is a unique global optimal contract combination.

Fine grid local search : take a smaller range around the b_L^c and b_H^c obtained by the coarse search, δ (for example 0.05), then the grid search can be carried out with a finer step size Δ_2 (for example 0.001) in the region, and the corresponding $b_L \in [b_L^c-\delta, b_L^c+\delta], b_H \in [b_H^c-\delta, b_H^c+\delta]$, and then calculate the EP of each grid one by one and find the maximum value (b_L,b_H). At this time, the selected area and step size range will be much smaller than before, so the points determined in this

interval can be used as the position of more accurate target points.

Figure 4: Heat map of shareholders ' expected return EP with respect to (b_L,b_H) under coarse grid

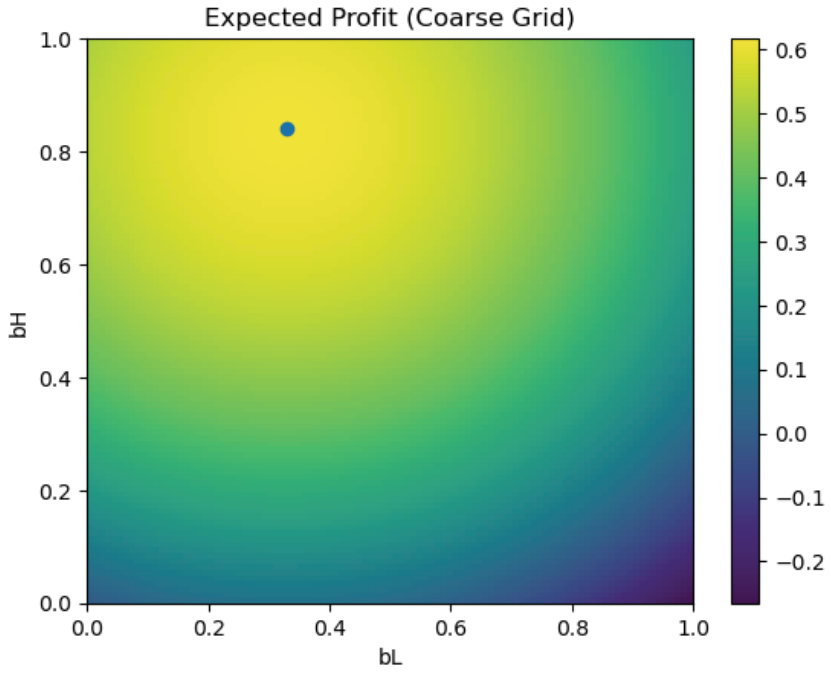


Figure 5: Local heat map of shareholders ' expected return EP under fine grid conditions

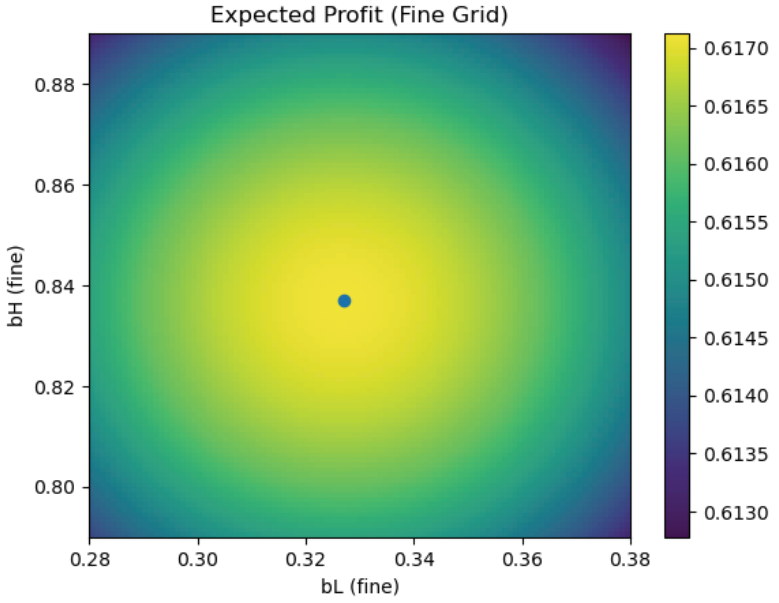
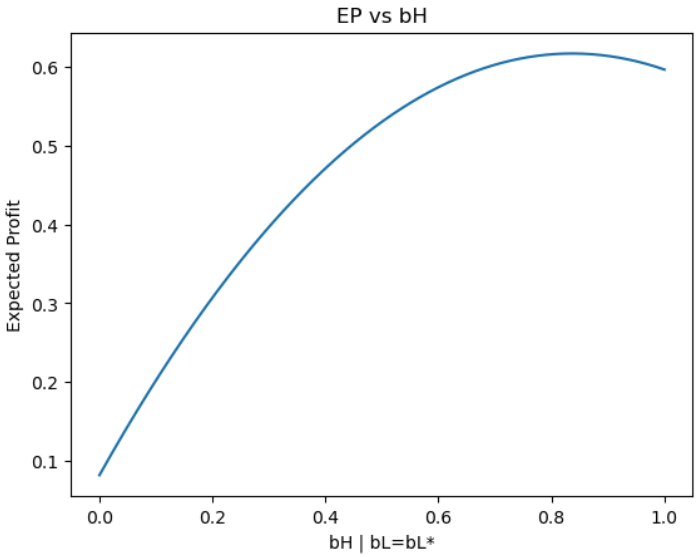


Figure 5 is the local heat map of the shareholder 's expected return EP under the fine grid condition (the parameter range has been scaled to the vicinity of the coarse optimal point), and the blue point represents the optimal point obtained by the fine search (b_L,b_H). In this local range, EP is a smooth unimodal function, so the peak point at the optimal point can be found by fine search.

After the above steps, the global optimal contract incentive coefficient combination (b_L,b_H) is obtained, and then the corresponding (w_L,w_H), (e_L,e_H) and all derivatives are calculated by substituting 3.4 binding constraints. This algorithm ensures that the optimal solution that satisfies the constraint conditions can be finally found, and it can effectively avoid falling into the local extremum. As long as there is a real global optimal value point in the fine search area, it is only necessary to ensure that δ is selected small enough, and at the same time ensure that there is sufficient margin near several coarse advantages. This study has verified the correctness of this method in the above way. On this basis, this study can construct a

cluster of numerical relationship curves between different factors based on different benchmark situations by adjusting the parameters.

Figure 6: The change of shareholders ' expected return EP with b_H (b_L,b_H)



As shown in Figure 6, when $b_L = b_L^*$ is fixed, the shareholder 's expected return EP changes with the change of b_H, that is, when $b_H = b_H^*$ (vertical line), the EP is the largest ; on the contrary, once b_H deviates from b_H^*, EP will drop sharply, indicating that shareholder returns are highly sensitive to b_H near the optimal high-capacity incentive intensity point, and the corresponding curve is also unimodal.

Figure 7: The cross-sectional view of the change of the shareholder 's expected return EP with b_L when $b_H=b_H^*$ is fixed

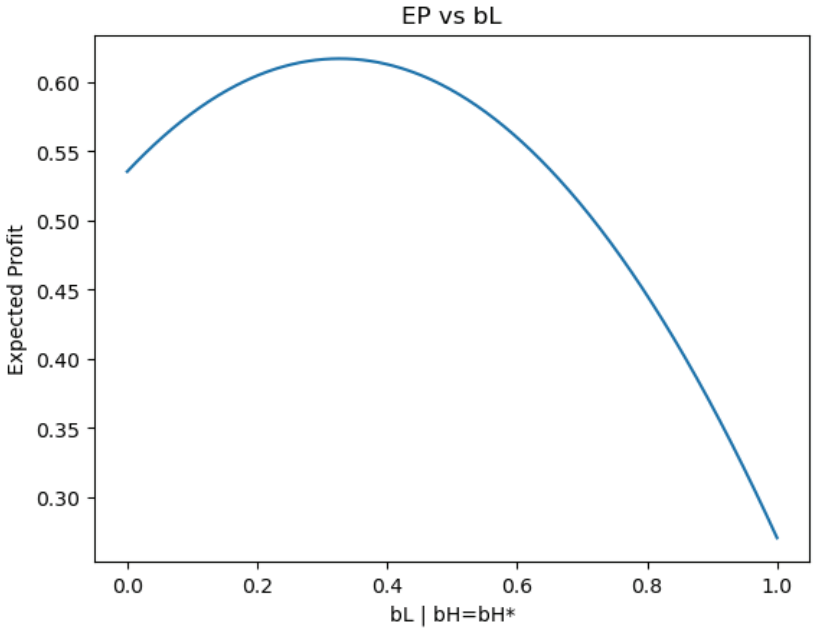


Figure 7 is a cross-sectional view of the change of the shareholder 's expected return EP with b_L when $b_H=b_H^*$ is fixed. Similarly, when $b_L=b_L^*$, the profit is the largest, and the deviation from the optimal point will lead to a decrease in EP. This shows that the optimal low-capacity incentive coefficient is also the only peak point of the income function. Combined with Figure 5, it can be seen that the second derivative of EP to parameter changes near (b_L^*,b_H^*) is negative, and the gain function at the optimal point satisfies a certain concavity.

Through the cross-sectional view, it can be seen more intuitively that the optimal contract parameter is the ' mountain top ' of the revenue function. This means that the (b_L^*,b_H^*) found in this study is indeed globally optimal, and shareholder returns are

quite sensitive to the deviations of these two parameters, which need to be accurately adjusted to achieve maximum. This also confirms the importance of designing differentiated contracts for different types of capabilities : if the incentive coefficient of high-capacity managers is set too low, the company 's earnings cannot fully tap its potential ; if the incentive coefficient of low-capacity managers is too high, the company 's income will also be damaged by ineffective incentives, and the optimal contract can balance the two.

4.3 Expandable Design

The solution path of the above model has good scalability and can be extended to different contract design situations. First, at the parameter level, the sensitivity of the model to some main parameters can be considered. For example, by changing the probability p of high and low ability types, the influence of the proportion of low ability of different proportion of manager groups on the optimal contract is simulated. Studies have shown that when the proportion of low-ability managers is greater, the value of distinguishing incentives is greater, and there are two kinds of contract design ; when the proportion of high ability is larger, shareholders can choose to give only a single type of contract to high-ability managers, that is, to carry out ' high-type ' incentives, without distorting the incentives of all managers for a very small number of low-ability managers. This study can test whether this statement is correct by adjusting the p, and find the condition value corresponding to the contract inflection point. Second, at the level of model construction, it can be extended from two types to continuous types. Although the solution of the continuous type requires more numerical algorithms, the existing model solving process in the two-type case can be extended by discretizing $b(\theta)$ into a finite number of parameters. Thirdly, other factors are added, such as long-term incentive, multi-task, agent fairness preference, etc. In this process, as long as the corresponding utility or constraint part is added to the model, the optimal contract combination can still be found by numerical method, and the solution algorithm of this study is insensitive to the form of constraint conditions and objective function, so the solution algorithm is very general. Of course, as the model increases more complexity, it will bring higher computational requirements, and the scheme proposed in this study can also further use more advanced optimization algorithms such as genetic algorithm and simulated annealing to assist it in searching for better optimal solutions. The scheme proposed in this study is actually an affordable framework for analyzing complex contract design problems. When it cannot be directly analyzed, the results can still be obtained by calculation.

5. Baseline Results and Policy Implications

5.1 Evaluation Caliber and Comparison Benchmark

In order to clearly interpret the optimal compensation contract, this study first clarifies the evaluation criteria and comparison criteria. In terms of evaluation caliber, shareholders are concerned about the company 's expected profit EP and its changes relative to different contractual arrangements. The CEO is concerned about their own utility, and the realization of utility is mainly affected by participation constraints and information rent. In the model of this study, because the utility is used to determine the equivalent form and the participation constraint is set to zero utility, if the CEO is low, its utility is ultimately 0 ; if it is high capacity, its utility is the size of information rent (IR(H) relaxation). Therefore, this study focuses on the excess utility of high-capacity CEOs and the difference in the effort level of the two types of CEOs. In terms of comparative benchmarks, there are two cases with reference significance : one is the complete information contract, that is, shareholders know the type of CEO ability in advance, and can formulate contracts for different types without information constraints. In this case, shareholders do not need to give high-capacity information rents, and incentives that meet the first-order optimal conditions can be set for each type ; the second is a single contract (regardless of type), that is, shareholders do not distinguish between CEO types and provide a unified contract for both types of people to accept. This is equivalent to giving up the use of adverse selection mechanism, only to solve the problem of moral hazard. These two benchmarks represent the ideal case under information symmetry and the sub-optimal case without separation mechanism, respectively. This study will compare the optimal contract with these two benchmarks when discussing the results to highlight the impact of information asymmetry on contract distortions. Since the complete information contract can be regarded as an extreme case of p taking 0 or 1, and the single contract is equivalent to the constraint condition of forcing $b_L=b_H$, the effect of the two can be observed by adjusting the model parameters and constraints. It should be noted that the solution of this study focuses on the separation contract,

because under reasonable parameters, when the two types coexist and p is moderate, the separation contract can bring higher shareholder returns. However, in extreme cases (such as high capacity probability is very large or very small), a single contract may become a sub-optimal choice for shareholders. This study will discuss this policy implication in Section 5.3.

5.2 Baseline Solution Results

Based on the baseline parameters (θ_L=1.0, θ_H=1.6, p=0.5, k=1.0, σ=1.0, γ=0.5, U_0=0) proposed above, the optimal contract parameters and corresponding performance indicators are obtained. The main results are summarized as follows: b_L^*=0.327, b_H^*=0.837, w_L^*=-0.027, w_H^*=-0.638 (Amount units can be understood as a standard currency, such as millions of dollars). The CEO 's effort level is calculated as $e_L^*=\frac{b_L^*\theta_L}{k}$=0.327, $e_H^*=\frac{b_H^*\theta_H}{k}$=1.3392. The expected profit obtained by shareholders is π_L^*=0.247, π_H^*=0.987, and the comprehensive expected return $EP^*\approx0.617$ (the same unit as above, which can be understood as million yuan). In addition, the verification of the constraint conditions shows that the relaxation of the high-capacity CEO 's participation in the constraint IR_H=0.0834, which is the information rent it obtains ; the incentive compatibility relaxation of low-capacity CEOs IC_L=0.463, indicating that the utility of high-capacity contracts is much lower than that of staying in their own contracts, proving that low-type IC constraints are stress-free, which is in line with the previous judgment of this study on constraint binding.

Figure 8: The elements of the optimal contract and the corresponding effort level histogram

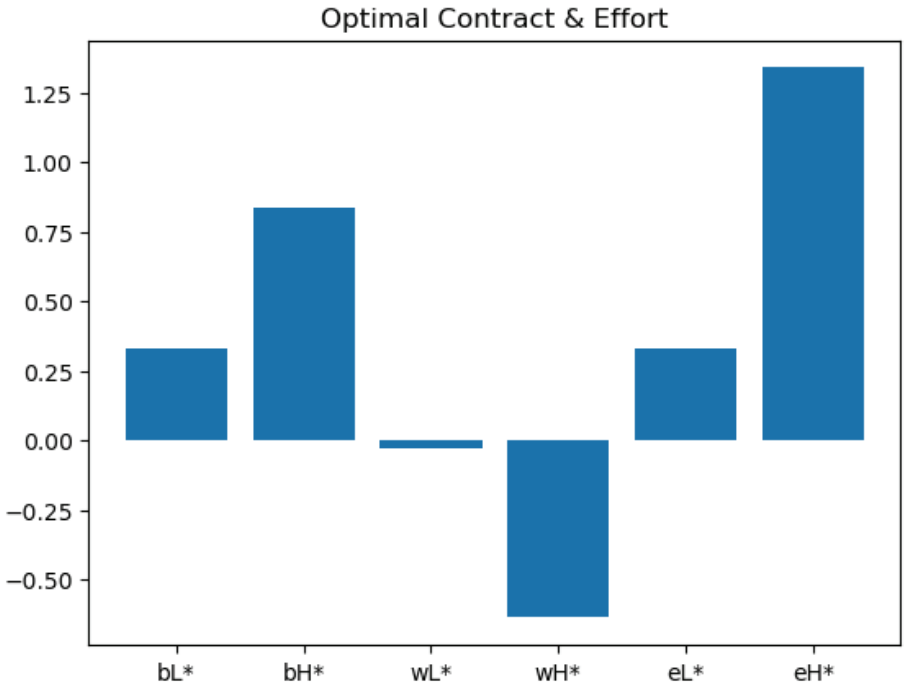


Figure 8 is a histogram of the elements of the optimal contract and the corresponding effort level. It can be seen from the figure that the optimal bonus coefficient b_L^*,b_H^*, fixed salary w_L^*,w_H^* and effort level e_L^*,e_H^* corresponding to low / high-capacity CEOs, the bonus coefficient of high-capacity CEOs is much higher than that of low-capacity CEOs, and the fixed salary is lower than that of low-capacity CEOs, and is negative, that is, high-capacity CEOs are required to pay a certain deposit or make corresponding investments after the contract is signed. In this way, the scheme of ' substantial high reward and low guarantee ' has the characteristics of stronger incentive and risk sharing for employees with stronger ability. On the contrary, the performance sensitivity of CEOs with low ability is low, and the fixed salary is only slightly less than 0 (if it is a small negative number, it requires a small amount of investment). There is also a big difference in the degree of effort between the two types of CEOs : the high-ability CEO 's effort is about 4.1 times that of the low-ability CEO, that is, under the design of the optimal contract, the high-ability CEO will be more motivated to work hard and create more performance.

From the above results, it can be seen that the optimal compensation contract adopts different incentives for different types of capabilities : high-capacity CEOs bear greater performance risks and obtain income mainly through performance, while low-capacity CEOs bear less performance risks and their income is more guaranteed by fixed parts. Specifically, $b_H^*\approx0.837$

means that for every unit of profit generated by a high-capacity CEO, 0.837 units can be extracted as bonuses, which is almost a ' high commission ' incentive ; correspondingly, $1-b_H^* \approx 0.163$ means that shareholders only retain 16.3% of their profits, but because of the high output efficiency of high-capacity CEOs, shareholders can still profit from the larger base brought by their efforts. $b_L^* \approx 0.327$ gives a moderate incentive intensity to low-capacity CEOs to share about 30% of marginal profits, and $1-b_L^* \approx 0.673$ guarantees shareholders ' large share of profits in low-capacity sectors. It is worth noting that both w_H^* and w_L^* are negative, especially the fixed salary $w_H^* \approx -0.638$ absolute value of high-ability CEO is larger. In theory, this means that high-capacity CEOs need to pay a certain amount (such as buying company shares or paying performance bonds) at the time of signing the contract to qualify for high performance bonuses. $w_L^* \approx -0.027$ is close to 0, indicating that low-capacity CEOs can hardly get a fixed salary, but they do not need to pay as much as high-capacity ones. This negative fixed wage may be understood in real contracts as executives voluntarily deferred part of their compensation or put it into the company to show confidence and bound interests. In fact, there are indeed arrangements for senior executives to hold company stocks, options, etc., which can be regarded as a disguised negative wage (senior executives use cash to purchase company equity, and future earnings depend entirely on the quality of operation). The results of this study theoretically support the necessity of this approach for high-capacity managers - by requiring their own deposit or shareholding, high-capacity managers can be fully motivated and their information rents are partially grabbed by shareholders.

5.3 Mechanism Interpretation and Governance Enlightenment

The above characteristics of the optimal contract are derived from the trade-off mechanism between incentives and rents under information asymmetry. In terms of high-capacity CEOs, in order to encourage their efforts, the contract gives a high proportion of revenue sharing b_H^*, which enables them to obtain a good return for every additional effort, so that they are willing to invest in the effort level of those with far ultra-low capabilities (1.3392 vs.0.327). At the same time, because of the strong ability, the high-ability CEO can easily overfulfill the low-ability contract and obtain more efficient use. In order to prevent it from doing so, shareholders must reduce the risk compensation of high-capacity contracts, that is, reduce the fixed salary w_H, until the high-capacity CEO does not benefit from the impersonation. The solution of this study shows that this process even puts the w_H value at a negative value, that is, letting the high-ability CEO 's own ' out of pocket ' to indicate their identity. At this time, there is no difference in the utility of high-capacity CEOs between the two contracts, and their information rent is just compressed to the extent that it is almost unprofitable to defraud low-capacity contracts (this model calculates that the high-capacity CEO still has about 0.083 utility excess, but it is very small relative to its effort output). On the other hand, for low-capacity CEOs, the contract provides a smaller performance incentive b_L^* has a double meaning: first, considering that the marginal output of low-capacity CEOs is not high, too strong incentives cannot effectively improve performance, but increase their risk burden ; second, more importantly, reducing b_L makes it difficult for high-capacity people to obtain higher returns than in high-capacity contracts because of the small proportion of bonuses, even if they consider posing as low-capacity to take such contracts. Similarly, a slightly higher fixed salary (or a smaller negative w_L) for low-capacity people can increase the attractiveness of low-capacity contracts to high-capacity people, but that would violate IC_H requirements. Therefore, shareholders also tend to set w_L at the lowest possible level when designing low-capacity contracts (until the lower limit of its IR_L is 0). The final result is that low-capacity contracts contain almost no rent and high incentives, so that high-capacity people have no incentive to fake, while low-capacity people can accept them themselves. As a result, two types of CEOs with different abilities are self-screened : high-ability CEOs choose high-incentive, high-risk and high-return contracts and work hard, while low-ability CEOs choose low-incentive and low-security contracts to fulfill their duties. Shareholders obtain the highest expected profit through this mechanism design, while minimizing the efficiency loss caused by information asymmetry.

The above mechanism has certain implications for the actual corporate governance and compensation design : First of all, the company should consider the differences in executive ability when formulating executive compensation, and cannot use the same structure of compensation to motivate everyone. The model of this study shows that a high proportion of performance bonuses should be provided to capable CEOs and fixed compensation can be reduced accordingly (even performance deposit or stock purchase requirements are set) to fully stimulate their talents and prevent efficiency losses caused by adverse

selection. This is a bit similar to the practice of ' high performance sensitivity contract ' or ' heavy performance, light bottom ' incentive scheme. On the contrary, for executives with relatively general capabilities, excessive performance sensitivity is not appropriate. Fixed compensation guarantees should be appropriately increased and performance appraisal efforts should be reduced so as not to give up or be negative because they cannot meet high indicators (this corresponds to low-capacity contracts in the model. Incentives are weak to ensure their participation). Such hierarchical incentives help to achieve pay for ability and do their best in the team. Secondly, the existence of information rent means that high-capacity executives tend to receive additional benefits beyond their retained wages, which may be reflected in the fact that excellent CEOs get generous equity incentives or bonuses. This is not detrimental to shareholders-it is precisely to pay a certain amount of information rent in exchange for executives to create greater value for the company. Therefore, when formulating executive compensation restriction policies, regulators should recognize that moderate executive excess returns under information asymmetry may be an incentive compatibility cost, and should not simply " one-size-fits-all " depress executive performance income, otherwise it may dampen the enthusiasm of high-ability people or drive them to conceal their talents. Finally, the analysis of this study also suggests that another strategy that companies can adopt when hiring executives is type screening : if they can screen out low-capacity candidates through assessment and only retain high-capacity CEOs when hiring, then p actually approaches 1. At this time, shareholders do not necessarily need to provide two types of contracts, and can directly give a unified strong incentive contract (equivalent to only designing contracts for high-capacity). Conversely, if the majority of low-capacity (p is significantly lower than 0.5), shareholders may also choose not to hire or eliminate high-capacity (because it is rare) and only design simple contracts for most low-capacity. However, in most cases, high and low abilities coexist and are difficult to be completely screened, so differentiated compensation contracts are still necessary. The above ideas reflect that the design of the optimal compensation contract should be combined with the ability structure of the executive group. This is consistent with the concept of formulating personalized incentive plans according to the characteristics of managers in management practice.

6. Conclusion and Prospect

6.1 Research Conclusions

This study studies the design of CEO 's compensation contract in the presence of moral hazard and adverse selection, and draws the following main conclusions:

(1)Type separation: Different CEOs are given different 'pay menu' salary schemes, so that they can decide what kind of contract to choose according to their own talents. High-energy CEOs will choose high-incentive-intensity contracts, while low-energy CEOs will choose the opposite, in order to solve the problem of adverse selection.

(2)Incentive intensity differentiation: The optimal contract increases the performance sensitivity and risk tolerance of high-capacity CEOs to the highest, and gives corresponding high returns, while for low-capacity CEOs, it intentionally lowers incentives. This strategic distortion is not the real optimal, but it is still efficient in the context of serious information asymmetry between the two parties.

(3)Compression of information rent: contract design minimizes the information rent of high-capacity CEOs, while low-capacity CEOs only get retained utility. This contract design minimizes the additional compensation that needs to be paid to executives due to information asymmetry.

(4)Improve shareholder returns: on the contrary, the ' one-size-fits-all ' single contract will reduce the expected return of shareholders compared to the separation contract, which indicates the importance of differentiated incentives, especially when the proportion of high-capacity and low-capacity executives is relatively balanced.

In conclusion, the study finds that the differentiated contract of ' high energy and high incentive, low energy and low guarantee ' can make executives self-screening and incentive compatibility, and maximize the interests of shareholders under the guarantee of minimum information rent.

6.2 Future Prospects

This study still has limitations and can be expanded from the following aspects in the future:

(1)Model hypothesis: We can consider extending the executive ability from two types to different continuous distribution types, and design the corresponding dynamic contract model on this basis to make it more in line with the actual situation.

(2)Governance environment: Since the model does not consider the contradiction between shareholders and the CEO of the board, if these corporate governance factors are taken into account, it will help to better explain the reality of incentive anomalies.

(3)Performance indicators: In the future, we can explore how to design the optimal incentive contract with multi-dimensional (not only financial) indicators under the dual information asymmetry.

(4)Empirical test: the data of listed companies can be used to empirically verify the core predictions of the model (such as the high-capacity CEO prefers high-incentive intensity contracts and their impact on shareholder returns).

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Wang, X., Zhang, L., Guan, X., Xiao, L., & Qian, G. (2016). Effects of Agent Fairness Sensitivity on Contracts in Moral Hazard Models. Journal of Management Sciences in China, 19(08), 21-31.

[2] Zhang, X., & Wang, Q. (2006). Research on the Optimal Incentive Contract Model for Venture Capitalists: A Game Theoretical Analysis under Adverse Selection and Moral Hazard. Journal of Finance and Economics, (05), 129-135. https://doi.org/10.16538/j.cnki.jfe.2006.05.012.

[3] Ni, X., Zhao, H., Huang, S., & Qian, L. (2020). Research on Compensation Contracts under Differences in Risk Control Capability. Chinese Journal of Management Science, 28(09), 23-32. https://doi.org/10.16381/j.cnki.issn1003-207x.2020.09.003.

[4] Ding, Y. (2003). Incentive and Restraint Mechanisms Considering Both Hidden Information and Hidden Action. The Journal of Quantitative & Technical Economics, (04), 122-125. https://doi.org/10.13653/j.cnki.jqte.2003.04.027.

[5] Zhang, W. (2005). Contract Linearization and Linearized Contracts. Chinese Journal of Management Science, (03), 62-67. https://doi.org/10.16381/j.cnki.issn1003-207x.2005.03.011.

[6] Mao, S. (2021). Compensation Contract Problems with Fuzzy Uncertainty under the Coexistence of Moral Hazard and Adverse Selection [Doctoral dissertation, Nanjing University of Science and Technology]. DOI:10.27241/d.cnki.gnjgu.2021.001788.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

A Study on the Spatio-Temporal Evolution and Determinants of Green Development Efficiency in the Cultural Industry of the Yangtze River Economic Belt

Can Qin*, Xuemei Zhang, Xiyue Yang, Rui Zhou, Yutong Cai

Chongqing Finance and Economics College, School of Economics, Chongqing, 401320, China

Corresponding author: Can Qin, 8834835478@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This study builds a super-efficiency SBM framework that takes unintended outputs into consideration. Using the Dagum Gini index, Moran's I statistic, and a principal component regression model, we evaluate the green development efficiency of the cultural sector across 11 provincial-level units within the Yangtze River Economic Belt for the year 2023. The analysis uncovers both the spatial-temporal dynamics and the driving forces behind efficiency. Our key findings are as follows. The overall green development efficiency of the cultural industry in this region remains fairly low, displaying a distinctive pattern: high efficiency in the lower reaches, weak technical capacity in the middle reaches, and limited scale in the upper reaches. Spatial differences mainly come from the gap between the lower and upper reaches, while significant internal variation exists within the upper reaches. Spatial clustering shows a clear "high-high" and "low-low" polarization with pronounced local agglomeration. The scale of the cultural industry acts as a positive driver, whereas energy intensity imposes a notable constraint. Against the background of the 15th Five-Year Plan, this paper puts forward a coordinated approach that spans policy alignment, business model transformation, technological empowerment, and talent development. The aim is to foster sustained improvement in green development efficiency and contribute to high-quality growth of the cultural industry.

Keywords: Yangtze River Economic Belt; Cultural Industry Efficiency; Green Development; Spatio-Temporal Evolution; Influencing Factors

Published: May 24, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1378

1. Introduction and Literature Review

1.1 Introduction

The "Opinions of the Central Committee of the Communist Party of China and the State Council on Accelerating the Comprehensive Green Transformation of Economic and Social Development" explicitly stipulate: "Promote the high-quality development of the cultural industry and foster the deep integration of culture and tourism. Actively encourage the accelerated development of new industries, new business models and new commercial formats oriented towards green and low-carbon development." [1] The "Guiding Opinions on Further Strengthening Ecological and Cultural Development" further propose: "Support the development of the 'ecological culture+' industry, explore the deep integration of ecological literary and artistic creation with science and technology, and actively utilise environmentally friendly materials to develop ecological cultural and creative products. " [2] Against this backdrop, the green development of the cultural industry has become an

integral part of implementing the “dual carbon” goals and driving the comprehensive green transformation of the economy and society. Its value lies in: guiding the low-carbonisation of the entire chain of cultural production and consumption through green concepts; fostering new growth poles for green consumption; and helping to establish green modes of production and lifestyles. At the same time, the “Outline of the 15th Five-Year Plan for National Economic and Social Development of the People’s Republic of China” emphasises, from a regional development perspective, that “we must adhere to the principle of prioritising major conservation efforts over large-scale development, uphold the principle of ecological priority and green development, and continuously advance the high-quality development of the Yangtze River Economic Belt.” [3] The Yangtze River Economic Belt is not only a priority area for national ecological conservation but also a region rich in cultural industry resources and characterised by dynamic development; exploring the spatio-temporal evolution of the green efficiency of the cultural industry in this region holds significant exemplary value. Consequently, this study takes the 11 provinces and municipalities of the Yangtze River Economic Belt as its research subjects. By considering a Super-Efficiency SBM model that accounts for unplanned outputs, and utilising the Dagum Gini coefficient, Moran’s index and principal component regression models, it systematically evaluates the spatio-temporal evolution and influencing factors of green development efficiency in the cultural industry. This provides empirical support for enhancing regional coordination and promoting the green transformation of the cultural industry, thereby contributing to the realisation of the goal of ecological and sustainable green development in the cultural sector.

1.2 Literature Review

When we look at the theoretical background, scholars have approached this topic from several different angles. Taking an enterprise perspective, Shi Xiaolong et al. (2025) pointed out a clear non-linear relationship between agglomeration and green transformation [4]. Other researchers like Chen Li (2025), Luo Lan (2025), and Huang Xin alongside Hu Angang (2025) shifted the focus slightly to examine the logic of transformation driven by new-quality productive forces [5][6][7]. Adding to this, Chen Yiping (2024) broke down the scientific connotations and structural elements of these new-quality productive forces specifically within the cultural sector [8]. All these studies basically lay a solid conceptual foundation that our own empirical research can build on.

Moving on to research methodology, there are a few standout approaches. Wan Qian (2025) actually used a super-efficiency SBM model to figure out the efficiency of the cultural industry right here in the Yangtze River Economic Belt, successfully tracking down the key influencing factors [9]. For spatial analysis, Li Jian et al. (2025) brought in the Dagum Gini coefficient and spatial differentiation methods. They used these specific tools to map out the spatio-temporal characteristics of green development efficiency for resource-based cities in the region [10]. On the policy side, Chen Lei et al. (2025) designed a quasi-natural experiment to measure the actual causal effects of regional economic policies on green technology innovation performance [11]. Another interesting angle came from Jiang Ziran et al. (2025), who mixed ecological efficiency metrics with strategic evaluation methods [12]. This helped prove how the overall development strategy of the Yangtze River Economic Belt boosts ecological efficiency along the entire corridor. Bringing these various methods into the mix has really made efficiency evaluation and policy identification a lot more precise.

Focusing directly on the Yangtze River Economic Belt, Zhang Xueliang et al. (2025) put together a comprehensive summary covering the outcomes of the implementation of the development strategy. They also suggested some key measures to keep in mind for the upcoming 15th Five-Year Plan period [13]. Technology's role is huge here, Yuan Liang et al. (2025) noticed that higher levels of digitalisation and intelligentisation create a major positive impact on green total factor productivity [14]. Looking at regional influences, Shan Baoyan et al. (2025) used Shandong Province to show how the cultural industry brings about noticeable spatial spillover effects on regional green development [15]. In terms of digital growth, Hu Yanglin (2025) dug into the intrinsic logic and practical pathways that allow new-quality productive forces to empower the digital cultural industry [16]. Other scholars mapped out broader strategies. Li Jinge (2024) relied on the new development philosophy to suggest practical pathways for high-quality development within the cultural industry [17]. Qiu Guiqie (2024) gave us a broad review centered around green cultural development [18], and Zhou Jianxin alongside Zhu Xueping (2025) went ahead and systematically tracked the annual progress in research concerning China's cultural industry as a whole [19].

Even with all this background work, the existing research definitely leaves some gaps. We are still missing a truly systematic measurement of the efficiency of green development in the cultural industry, especially one that includes a holistic evaluation of river basins. Beyond that, past studies haven't done enough to unpack the spatio-temporal evolution of efficiency or pinpoint the specific regional differentiation characteristics. We also really need a deeper dive into the non-linear effects of various influencing factors. To try and fix this, our paper runs a super-efficiency SBM model that specifically accounts for unplanned outputs. We paired this up with the Dagum Gini coefficient, Moran's index, and multiple linear regression models. Doing this lets us carry out a thorough, scientific evaluation of the green development efficiency of the cultural industry across the Yangtze River Economic Belt. By clearly mapping out its spatio-temporal characteristics and influencing factors, we hope to build a really solid foundation that helps achieve those big strategic objectives of high-quality development in the cultural industry.

2. Model Construction, Method Selection and Indicator System Development

2.1 Model Construction

To scientifically evaluate the green development efficiency of the cultural industry in the Yangtze River Economic Belt and identify its influencing factors, this paper constructs two main types of models: firstly, a super-efficiency SBM model that accounts for unwanted outputs for efficiency measurement; and secondly, a principal component regression model for identifying influencing factors.

2.1.1 The Super-Efficiency SBM Model Accounting for Unintended Outputs

The Super-Efficiency SBM model, with SBM standing for Slacks Based Measure of Super Efficiency, provides a way to evaluate efficiency that is neither radial nor angular. Its calculation is built on slack variables. By taking into account undesirable outputs (such as pollutants), this model permits efficiency values greater than 1 for efficient decision-making units and addresses slack issues in inputs and outputs, thereby enabling a more precise evaluation of the green development efficiency of the cultural industry. The core formula is as follows:

$$\rho^{*}=\min\frac{\frac{1}{m}\sum_{i=1}^{m}\left(\frac{\overline{x}}{x_{ik}}\right)}{\frac{1}{r_1+r_2}\times\left(\frac{\sum_{s=1}^{r_1}\overline{y^{d}}}{y_{sk}^{d}}+\frac{\sum_{q=1}^{r_2}\overline{y^{u}}}{y_{qk}^{u}}\right)} \tag{1}$$

Where $\rho^{*}$is the calculated green development efficiency value; m is the number of input indicators; r_1 and r_2 are the numbers of desired and undesired outputs, respectively; x_{ik} is the i-th input; y_{sk}^{d} and y_{qk}^{u} are the desired and undesired outputs, respectively; $\overline{x}$, $\overline{y}^{d}$, and $\overline{y}^{u}$ are the input-output values adjusted for slack variables.

2.1.2 Principal Component Regression Model

To identify the key factors influencing the efficiency of green development in the cultural industry, a principal component analysis (PCA) is first performed on the three variables with the highest correlation to extract the first principal component (PC1). PC1 is then used as an explanatory variable in a regression analysis with technical efficiency (TE). The core formula for principal component regression is:

$$\mathrm{PC1}=\alpha_1\mathrm{urban}+\alpha_2\mathrm{cult}_{\mathrm{share}}+\alpha_3 X_2+\beta_3\mathrm{energy}_{\mathrm{int}} \tag{2}$$

where α is the loading (eigenvector) obtained from the principal component analysis; β is the regression coefficient; urban is the urbanisation rate; cult_share is the share of the cultural industry's value added in GDP; and energy_int is the energy consumption intensity.

2.2 Method Selection

To get a scientific measure of the green development efficiency of the cultural industries in the Yangtze River Economic Belt and how it evolves over time and space, this study brings together several quantitative methods. The Analytic Hierarchy Process and the entropy method are applied to assign weights to composite indicators, including the environmental pollution index and the cultural product consumption satisfaction index, and then to work out the composite scores. In addition, the Dagum Gini coefficient and Moran's I index are used to uncover spatial differences in efficiency and to identify patterns of agglomeration.

2.2.1 Indicator Weighting Methods

Given that the Super-Efficiency SBM model, which accounts for unintended outputs, requires the estimation of indices for "environmental pollution" and "consumer satisfaction with cultural products and services," the Analytic Hierarchy Process (AHP) and the entropy method were employed to determine the weights.

2.2.1.1 Analytic Hierarchy Process

A pairwise comparison matrix $B=(b_{ij})_{p\times p}$ is constructed, where b_{ij} represents the relative importance of indicator i to j(on a scale of 1 to 9). The core formula for calculating the weight vector is as follows:

Calculate the product of the elements in each row of the decision matrix M_I:

$$M_I=\prod_{j=1}^{P} b_{ij}\ (\text{i=1,2,…,p}) \tag{3}$$

Compute the pth root of M_i ω_i':

$$\omega_i'=\sqrt[p]{M_i} \tag{4}$$

Normalise to obtain the weight vector ω_i:

$$\omega_i=\frac{\omega_i'}{\sum_{j=1}^{p}\omega_j'} \tag{5}$$

Calculate the largest eigenvalue λ_{max}:

$$\lambda_{max}=\frac{1}{p}\sum_{i=1}^{p}\frac{(BW)}{\omega_i} \tag{6}$$

Consistency check:

$$\text{CI}=\frac{\lambda_{\max}}{\text{p-1}} \tag{7}$$

$$\text{CR}=\frac{\text{CI}}{\text{RI}} \tag{8}$$

2.2.1.2 Entropy Method

The entropy method is an objective weighting method based on information entropy. The calculation steps are as follows:

Normalisation of inverse indicators (environmental pollution indicators):

$$X'_{ij}=\frac{1}{X_{ij}} \tag{9}$$

Calculation of the weight b_{ij} for the jth indicator of the ith sample:

$$b_{ij}=\frac{X_{ij}}{\sum_{i=1}^{n}X_{ij}} \tag{10}$$

Calculate the entropy value e_j:

$$e_j=-K\sum_{i=1}^{n} b_{ij}\,I_n b_{ij} \tag{11}$$

Calculate the coefficient of variation d_j:

$$d_j=1-e_j \tag{12}$$

Calculate the weight ω_j:

$$\omega_j=\frac{d_j}{\sum_{i=1}^{p}d_i} \tag{13}$$

Final composite score:

$$S_i=\sum_{j=1}^{p}\omega_j X'_{ij} \tag{14}$$

2.2.2 Dagum Gini Coefficient

The Dagum Gini coefficient decomposes overall variance into within-group variance, net between-group variance and hypervariability. This method is suitable for multi-regional hierarchical analysis and can accurately identify the primary

sources of regional differences. This study applies the Dagum Gini coefficient to measure the disparities in green development efficiency of the cultural industry across the 11 provinces and municipalities in the Yangtze River Economic Belt. The analysis covers both the overall level and subregional comparisons. It also breaks down the contributions made by different groups, which helps identify where the main regional differences come from.

Let the population consist of k regions, with the jth region comprising n_j provinces and cities, and the total number of provinces and cities in the population being $n=\sum_{j=1}^{k} n_j$. Let y_{ji} denote the efficiency value of the ith province or city in the jth region, and $\bar{y}$ denote the population mean.

The overall Gini coefficient G is defined as:

$$G=\frac{\sum_{j}^{k}\sum_{h=1}^{k}\sum_{r=1}^{nj}\left|y_{ij}-y_{hr}\right|}{2n^{2}\bar{y}} \tag{15}$$

This can be decomposed as:

$$G=G_w+G_{nb}+G_t \tag{16}$$

where G_w represents the within-group variance contribution, reflecting the differences among provinces and cities within each region; G_{nb} represents the net between-group variance contribution, reflecting the differences in average levels between different regions, without considering overlap; and G_t represents the density contribution, reflecting the differences arising from sample overlap between different regions.

2.2.3 Moran's I

Moran's I is a statistical measure used to assess spatial autocorrelation, comprising the global Moran's I and the local Moran's I. This paper employs Moran's I to calculate the global and local spatial autocorrelation indices for the green development efficiency of the cultural industries across the 11 provinces and municipalities in the Yangtze River Economic Belt, and to identify the agglomeration types of each region, thereby revealing their spatial agglomeration patterns. The primary formula for Moran's I is as follows:

Formula for the global Moran's I:

$$I=\frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}(D_i-\bar{D})(D_j-\bar{D})}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\right)\sum_{i=1}^{n}(D_j-\bar{D})^{2}}(i\neq j) \tag{17}$$

where n is the number of provinces and cities; D_i and D_j are the efficiency values of provinces and cities i and j, respectively; $\bar{D}$ is the mean efficiency; and w_{ij} is an element of the spatial weight matrix. This study adopts a binary spatial weight matrix based on the Queen's adjacency rule: if provinces and cities i and j share a common boundary or vertex, then $w_{ij}=1$; otherwise, $w_{ij}=0$.

Formula for the local Moran's I:

$$I_i=\frac{n(D-\bar{D})\sum_{j=1}^{m}w_{ij}(D_j-\bar{D})}{\sum_{i=1}^{n}(D-\bar{D})^{2}}(i\neq j) \tag{18}$$

where m is the number of provinces and cities adjacent to province or city i.

2.3 Construction of the Indicator System

2.3.1 Basis for Indicator Selection

Based on the theory of endogenous economic growth, this paper selects input indicators from four aspects:labour, capital, technology and energy; it takes operating revenue from the cultural industry and consumer satisfaction with cultural products as desired outputs, and a composite index of industrial "three wastes" as undesired outputs. The study covers the 11 provinces and municipalities within the Yangtze River Economic Belt, with data primarily sourced from the 2024 National Statistical Yearbook; missing values have been imputed using interpolation.

2.3.2 Input-Output Indicator System

This paper employs four input indicators, two desired output indicators and one undesired output indicator, and utilises a super-efficiency SBM model that incorporates the undesired output to measure efficiency. The specific composition of the indicators and data sources are shown in Table 1.

Table 1 Efficiency Evaluation Indicators for Green Development in the Cultural Industry

ategory	Indicator Name	Indicator Composition	Unit	Indicator Explanation	Data Source
Input Indicators	Human Resources	Number of employees in the cultural industry	People	Personnel across cultural manufacturing, trade, and services, reflecting human capital investment scale.	*China Statistical Yearbook on Culture and Related Industries*
	Capital Investment	Assets of legal entities in the cultural and related industries	100 million yuan	Capital stock and production scale, the material foundation of cultural industry development	《China Statistical Yearbook on Culture and Related Industries》
	technology Investment	Internal expenditure on R&D	10,000 yuan	R&D investment intensity, reflecting cultural industry's innovation and green transformation potential	Statistical Yearbooks of Provinces and Municipalities
	Energy Input	Energy Consumption	10,000 tonnes	Energy input scale in cultural industry operations, a key constraint for green efficiency assessment	*China Energy Statistical Yearbook*
Output Indicators	Expected Output	Revenue of the cultural industry	10,000 yuan	Operating revenue-weighted index of cultural manufacturing, trade, and services, reflecting economic performance	*China Statistical Yearbook on Culture and Related Industries*
		Consumer Satisfaction with Cultural Products and Services	—	Weighted satisfaction index of cultural consumption, reflecting social benefits	Questionnaire survey
	Unintended Outcomes	Environmental pollution	—	Weighted index of industrial wastewater, air emissions, and solid waste, reflecting environmental externalities	Statistical Yearbooks of Provinces and Municipalities

2.3.3 Variable System for the Principal Component Regression Model

To investigate the factors influencing the efficiency of green development in the cultural industry, this study selects technical efficiency (TE) as the dependent variable, and per capita GDP (rgdp), urbanisation rate (urban), energy intensity (energy_int), the share of cultural industry value added in GDP (cult_share), and the urban-rural cultural consumption gap (cult_gap) as independent variables. As shown in Table 2.

Table 2 List of Variable Definitions

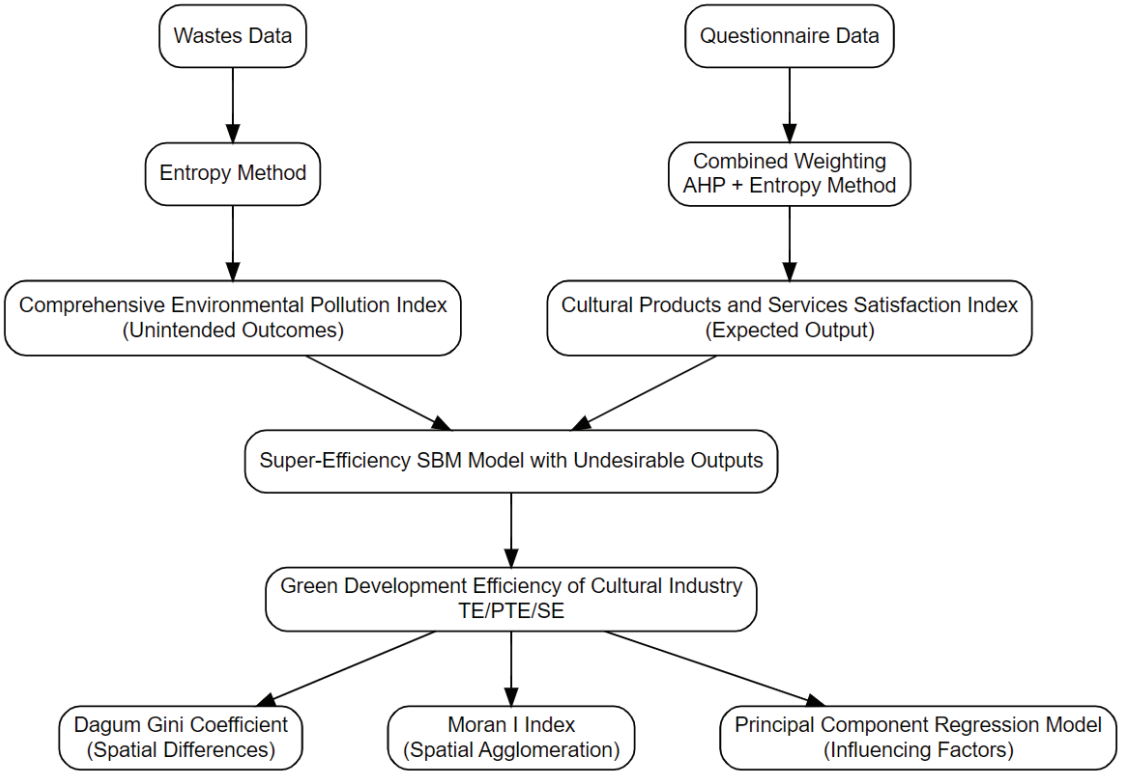
Variable	Variable Name	Unit	Explanation of Indicators
TE	Technical efficiency	-	Measuring the Green Development Efficiency of the Cultural Industry Using a Super-Efficient SBM Model that Accounts for Unwanted Outputs
rgdp	GDP per capita	100 million yuan	Reflecting the impact of economic development levels on green efficiency
urban	Urbanisation rate	%	Reflecting the impact of population concentration and urban development on efficiency
energy_int	Energy intensity	%	Measures energy efficiency and influences the sustainability of green development
cult_share	Share of cultural industry value added in GDP	%	Reflecting the impact of industry scale on green efficiency
cult_gap	Urban-rural cultural consumption gap	%	Reflecting the constraints on efficiency imposed by urban-rural consumption imbalances

It should be noted that, due to multicollinearity among the variables and the small sample size, both rgdp and cult_gap were left out of the final principal component regression model. They were not statistically significant in the univariate tests and also showed a strong correlation with the principal component. As a result, the model only included the first principal component (PC1), which was constructed from urban, cult_share, and energy_int.

3. Empirical Analysis

We went ahead and ran a comprehensive empirical analysis to see exactly how the green development efficiency of the cultural industry is holding up across the 11 provinces and municipalities in the Yangtze River Economic Belt. We started by estimating and breaking down the green development efficiency for each specific province and municipality. To get this right, we relied on the super-efficiency SBM model, making sure it accounts for unplanned outputs. After getting those numbers, we used the Dagum Gini coefficient along with Moran's I to pull back the curtain on spatial variations in efficiency and spot any real agglomeration characteristics. To wrap things up, we plugged the data into a multiple linear regression model to safely identify the key influencing factors driving these changes.

Figure 1 Flowchart of the empirical analysis approach



3.1 Analysis of Green Development Efficiency in the Cultural Industry Using a Super-Efficiency SBM Model Incorporating Unwanted Outputs

Building upon relevant literature, this paper constructs an evaluation index table for the green development efficiency of the cultural industry (Table 1) and selects the super-efficiency SBM model, which accounts for undesirable outputs, to measure the green development efficiency of the cultural industry across the 11 provinces and municipalities in the Yangtze River Economic Belt. Prior to running the model, it is necessary to calculate the environmental pollution index using the entropy method based on “three wastes” statistical data, and to calculate the cultural product and service consumption satisfaction index using a combination of the Analytic Hierarchy Process (AHP) and the entropy method based on questionnaire data.

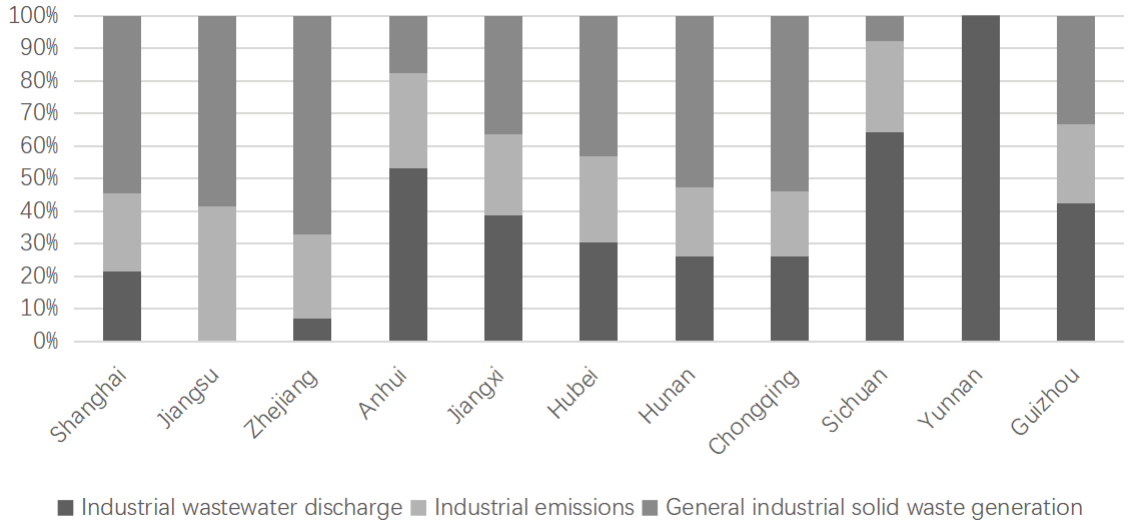
This paper employs the super-efficiency SBM model, which accounts for unwanted outputs, to measure the green development efficiency of the cultural industries in the 11 provinces and municipalities of the Yangtze River Economic Belt. An evaluation index table for the green development efficiency of the cultural industries (Table 1) has been constructed. Prior to running the model, it is necessary to calculate the environmental pollution index using the entropy method based on

statistical data on the "three wastes", and to calculate the consumer satisfaction index for cultural products and services using a combination of the Analytic Hierarchy Process (AHP) and the entropy method based on questionnaire data.

3.1.1 Calculation of the Environmental Pollution Index

Based on the statistical data on the "three wastes", each indicator was subjected to negative standardisation. Subsequently, the following were calculated for each indicator: the proportion of, the weight b_{ij}, the entropy value e_j, and the coefficient of variation d_j, yielding the weights ω_j. Finally, a weighted sum was taken to obtain the comprehensive environmental pollution index $EPI_i = \sum_{j=1}^{n} \omega_j x_{ij}^{'}$ for each region (as shown in Figure 2), serving as a non-desired output indicator.

Figure 2 Stacked bar chart of environmental pollution index percentages



3.1.2 Calculation of Satisfaction with the Consumption of Cultural Products and Services

Consumer satisfaction with cultural products and services, as an expected output indicator in the evaluation of green development efficiency, utilises basic data obtained through questionnaire surveys. The comprehensive satisfaction index is calculated by combining the Analytic Hierarchy Process (AHP) with the entropy method for weighting.

3.1.2.1 Sampling Design

A proportional to size probability sampling (PPS) method was adopted, with the 11 provinces and municipalities of the Yangtze River Economic Belt as the study area. One core city was selected from each province, using the 2023 permanent resident population as the size variable. Under a 95% confidence level and a 5% absolute error margin, the required sample size was calculated to be 502. A total of 500 questionnaires were distributed, with 458 valid responses returned, yielding a response rate of 91.6%.

3.1.2.1 Reliability and Validity Analysis

Table 3 Reliability Test for the Formal Survey and Validity Test

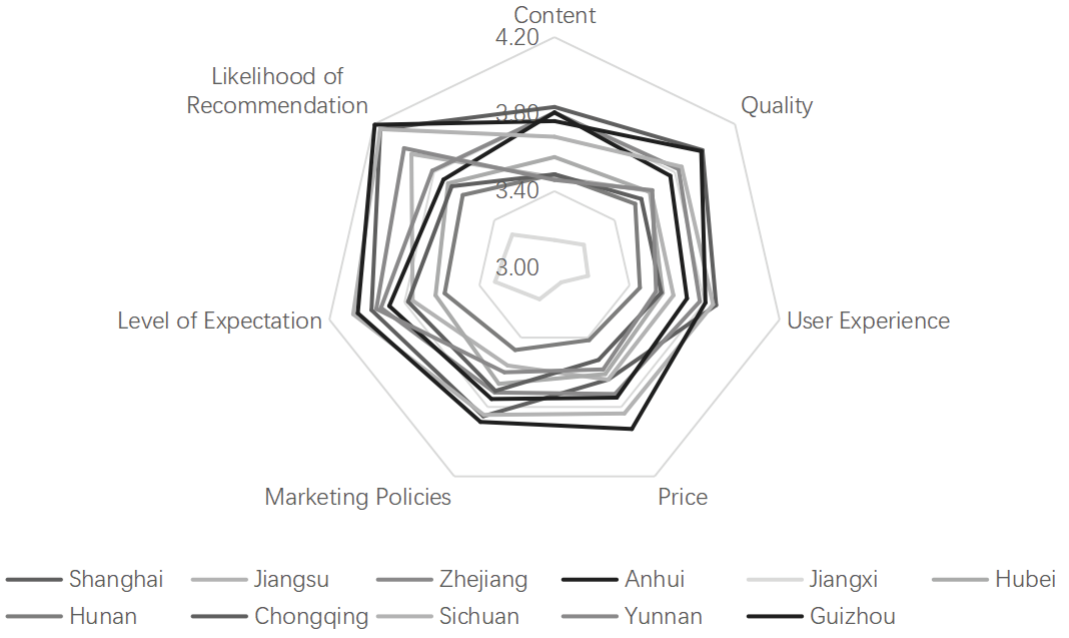
Item	CRONBACH'S ALPHA	Number of Items	Reliability Evaluation	KMO and Bartlett's test		
Q6	0.851	3	High	KMO()value		0.931
Q7	0.797	3	Fair			
Q8	0.701	3	Better	Bartlett's test of sphericity	Approximate chi-squared	1481.276
Q9	0.773	3	Fair			
Q10	0.749	3	Fair			
Q11&Q12	0.703	2	Fair		df	136
Overall scale	0.96	17	Very good		P-value	0

As shown in Table 3, the overall Cronbach's alpha coefficient for the questionnaire is 0.96, the KMO value is 0.931, and the p-value for Bartlett's sphericity test is<0.001; the questionnaire demonstrates good reliability and validity.

3.1.2.1 Calculation of Satisfaction with Cultural Products and Services

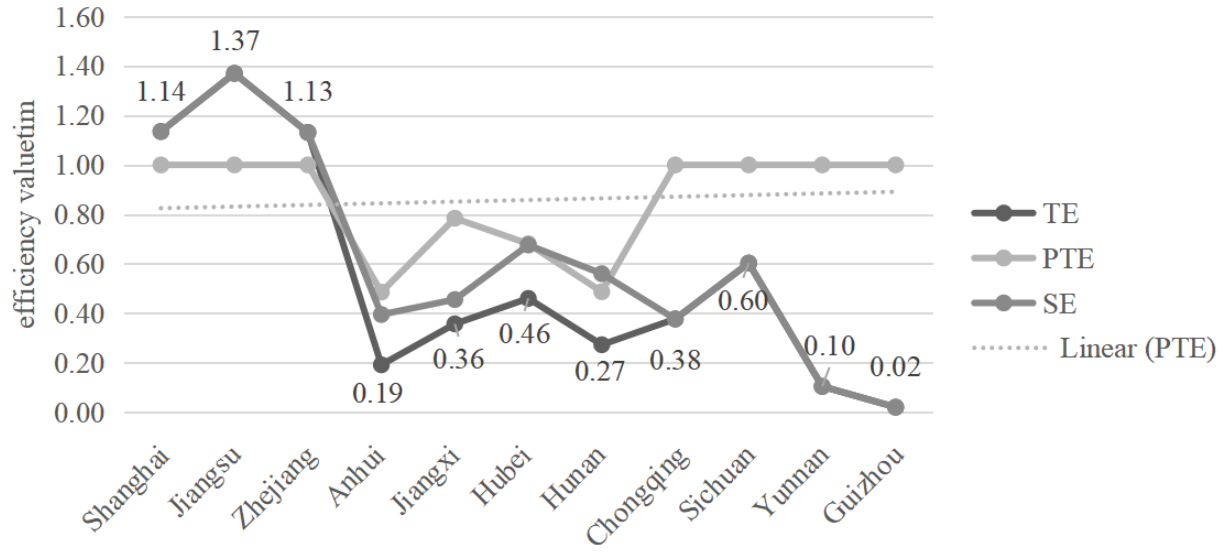
A combination of the Analytic Hierarchy Process (AHP) and the entropy method was employed for weighting. Firstly, the AHP was used to determine subjective weights: a satisfaction hierarchy (objective level, criterion level, and indicator level) was constructed; experts provided scores to create a judgment matrix, which passed the consistency test (CR < 0.1). Secondly, the entropy method was used to calculate objective weights based on data from 458 valid questionnaires. Finally, the mean of the AHP weights and the entropy-based weights was taken as the composite weight for each dimension. The weighted composite satisfaction score for each respondent was calculated, and these were aggregated by province and municipality to derive the average composite satisfaction index (Figure 3), which serves as an expected output variable in the super-efficiency SBM model that accounts for unanticipated outputs.

Figure 3 Radar chart of the cultural products and services consumption satisfaction index



3.1.3 Efficiency Measurement of the Super-Efficient SBM Model Considering Unwanted Outputs()

Figure 4 Line Chart of Efficiency Measurement



According to the calculation results (Figure 4), breaking the numbers down to look at pure technical efficiency (PTE), we saw seven different provinces and municipalities hit a perfect value of 1.0. This group included Shanghai, Jiangsu, Zhejiang, Chongqing, Sichuan, Yunnan, and Guizhou. Things weren't as great in the middle reaches of the Yangtze River, where Anhui (0.486), Jiangxi (0.784), Hubei (0.681), and Hunan (0.488) all struggled and fell below the 0.8 mark. Shifting the focus to scale efficiency (SE), only the three provinces located in the lower Yangtze River region actually managed to exceed 1. Meanwhile, every single region sitting in the middle and upper Yangtze River areas dropped below 0.7. In fact, if you just look at the four provinces making up the upper Yangtze River region—Chongqing, Sichuan, Yunnan, and Guizhou—their average SE was just a really low 0.277.

Looking at the map by region, the Lower Yangtze region leads across the board. They pulled in an average TE of 1.212, a perfect average PTE of 1.0, and an average SE of 1.212. The Middle Yangtze region tells a different story with an average TE of 0.321, an average PTE of 0.610, and an average SE of 0.522, which points to a pattern of dual insufficiency. Up in the Upper Yangtze region, we recorded a mean TE of 0.277 alongside a mean PTE of 1.0 and a mean SE of 0.277. This basically indicates that while they have pure technical efficiency, they are severely lacking in scale. Almost every non-efficient province and municipality ended up showing positive slack in unanticipated output, which usually means environmental pollution. On the flip side, the three provinces in the Lower Yangtze region exhibited negative slack when it came to capital input, suggesting some redundancy.

3.2 Spatial Differences Analysis Based on the Dagum Gini Coefficient

To reveal the sources of spatial disparities in the green development efficiency of the cultural industry within the Yangtze River Economic Belt, this paper employs the Dagum Gini coefficient and its decomposition method for analysis.

The Dagum Gini coefficient and its decomposition method were used to measure spatial variations in the green development efficiency of the cultural industry in the Yangtze River Economic Belt in 2023.

Figure 5 Bar chart of the Dagum Gini coefficient

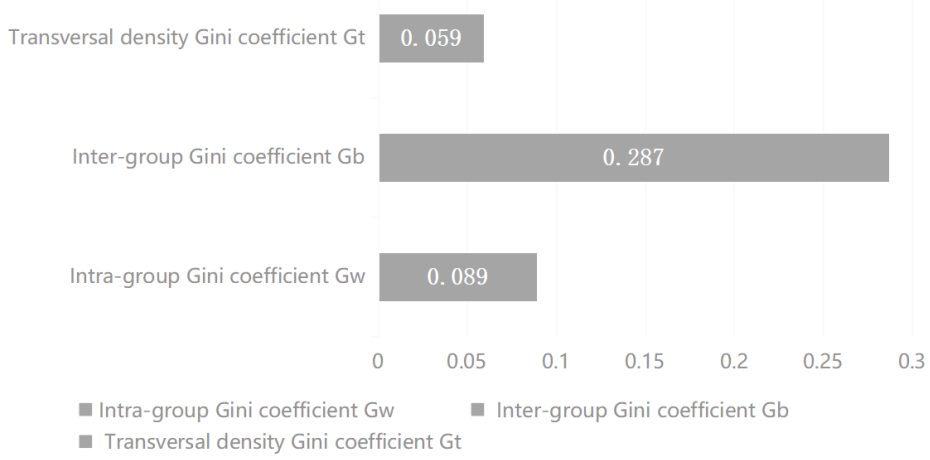
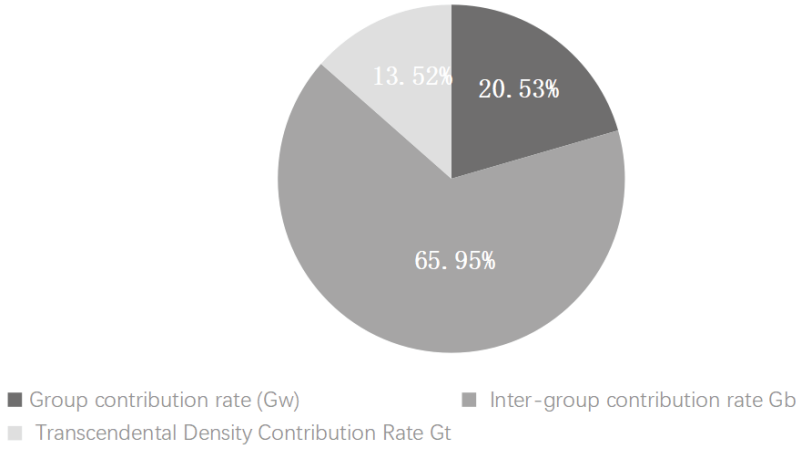
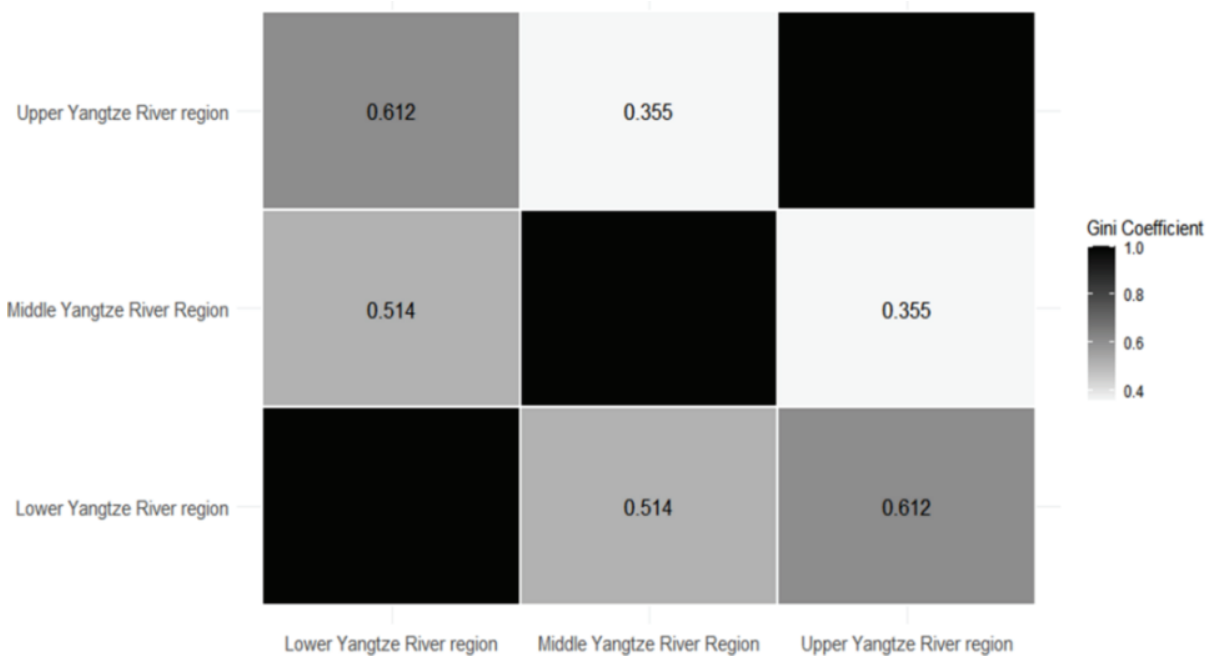


Figure 6 Bar chart of the Dagum Gini coefficient contribution rates



Looking at the actual calculations for the Dagum Gini coefficient shown in Figure 5, the overall Gini coefficient for the green development efficiency of the cultural industry across the Yangtze River Economic Belt sat at 0.436 for the year 2023. This number points straight to a pretty significant spatial imbalance between the different regions. When we break down where this disparity is coming from using the contribution rates in Figure 6, the inter-group contribution rate is clearly the main driver at 65.95%. This completely overshadows both the intra-group contribution rate, which is 20.53%, and the super-variable density contribution sitting at 13.52%. Now, if we zoom in on those intra-group differences, the Gini coefficient hits its peak within the Upper Yangtze region at 0.456. This really highlights some stark disparities in efficiency levels right among its constituent provinces and municipalities—namely Chongqing, Sichuan, Yunnan, and Guizhou. Trailing behind that is the Lower Yangtze region at 0.231, making the Middle Yangtze region the most internally balanced of the bunch with a coefficient of just 0.114. Further analyse the degree of variation within each region and between pairs of regions.

Figure 7 Dagum Gini coefficient variation heatmap



Based on the decomposition of the Dagum Gini coefficient from Figure 7, it is clear that these spatial disparities mostly originate from a stark contrast between the Lower Yangtze region and the Middle and Upper areas. While the Upper Yangtze region struggles with severe internal differentiation, the Middle Yangtze remains relatively homogeneous. As we noted, the internal Gini coefficient for the Upper Yangtze region is incredibly high at 0.456, which easily dwarfs the 0.231 of the Lower Yangtze and the 0.114 of the Middle Yangtze. This huge gap in the Upper region is driven by extremely severe efficiency divergence among its four provinces, especially when comparing Sichuan at 0.6031 with Guizhou way down at 0.0208. The Lower Yangtze region takes second place for internal disparity, largely because Jiangsu's score of 1.3699 is significantly higher than what we see in both Shanghai and Zhejiang. The Middle Yangtze region remains the most balanced internally, keeping its efficiency relatively concentrated across its own four provinces.

Shifting to an inter-regional perspective, the gap between the Lower and Upper Yangtze regions is by far the widest, showing a Gini coefficient of 0.612. This is followed by the gap between the Middle and Lower regions at 0.514, leaving the smallest measured gap between the Middle and Upper regions at just 0.355. Ultimately, this data demonstrates that the Lower Yangtze region leads comprehensively across the board in terms of efficiency. It creates a sharp, distinct divide with the Upper region, while the actual average gap between the Middle and Upper regions stays relatively small.

3.3 Analysis of Regional Aggregation in the Cultural Industry Based on the Moran Index

To reveal the spatial dependence and agglomeration patterns of green development in the cultural industries of the Yangtze River Economic Belt, this paper employs the Moran's I index to conduct a spatial autocorrelation analysis.

3.3.1 Analysis of the Moran's I Index

Table 4 Overall Moran's I

Study Item	Moran's I	Expected Value E(I)	Standard Deviation sd(I)	z-value	p-value
TE	0.527	-0.1	0.193	3.178	0.001

Table 5 Local Moran's Index

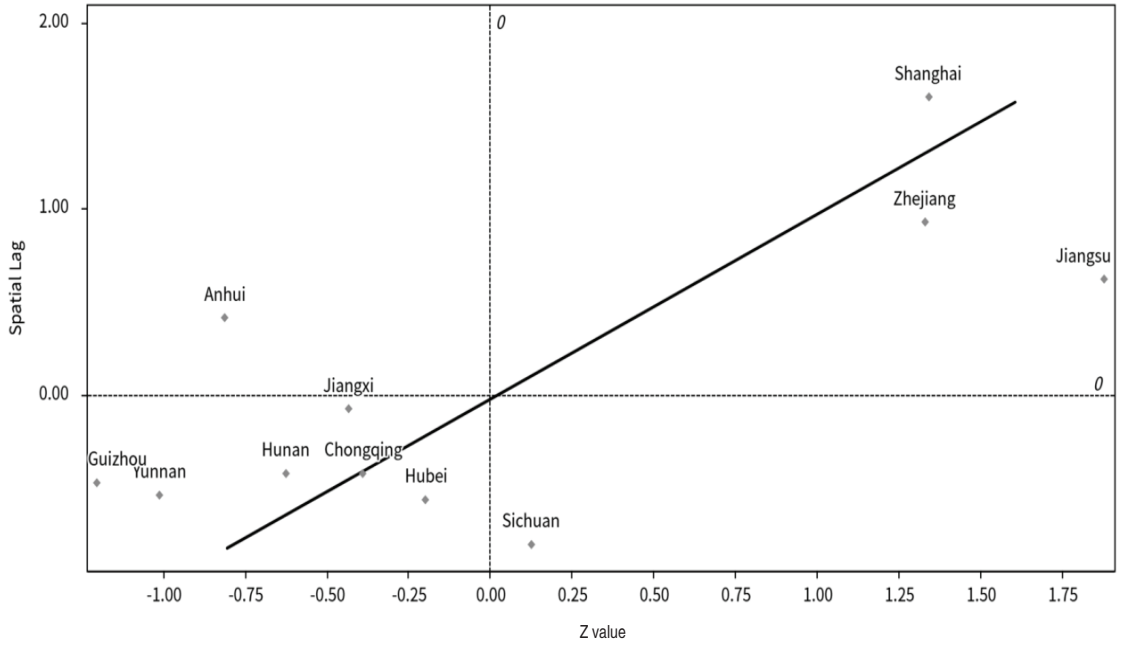
Location	Moran's I	Expected value E(I)	Standard deviation sd(I)	z-value	p-value
Shanghai	1.953	-0.1	0.767	2.866	0.001
Jiangsu Province	1.057	-0.1	0.83	1.66	0.024
Zhejiang Province	1.122	-0.1	0.661	1.929	0.013
Anhui Province	-0.306	-0.1	0.429	-0.533	0.149
Jiangxi Province	0.03	-0.1	0.189	0.221	0.206
Hubei Province	0.102	-0.1	0.087	1.197	0.058
Hunan Province	0.243	-0.1	0.264	1.044	0.074
Chongqing	0.15	-0.1	0.209	0.839	0.1
Sichuan Province	-0.093	-0.1	0.054	-1.705	0.022
Yunnan Province	0.494	-0.1	0.647	0.865	0.097
Guizhou Province	0.519	-0.1	0.472	1.385	0.042

The global Moran's I statistic (Table 4) shows I = 0.527, p=0.001, indicating that there is significant positive spatial autocorrelation in the green development of the cultural industry across the Yangtze River Economic Belt as a whole.

According to the local Moran's I results in Table 5, Shanghai (I = 1.953, p = 0.001), Jiangsu (I = 1.057, p = 0.024), Zhejiang (I = 1.122, p = 0.013), and Guizhou (I = 0.519, p = 0.042) each show a significantly positive spatial autocorrelation. Sichuan (I = −0.093, p = 0.022), on the other hand, has a significantly negative value. The remaining six regions, namely Anhui, Jiangxi, Hubei, Hunan, Chongqing, and Yunnan, are all statistically insignificant (p > 0.05 in each case).

3.3.2 Moran's scatter plot and quadrant distribution analysis

Figure 8 Moran scatter plot



Based on the quadrant distribution of the Moran scatter plot (Figure 8), the overall pattern shows high values in the east and low values in the west, with localised polarisation. Shanghai, Jiangsu and Zhejiang fall into the first quadrant (high-high), indicating a clustering of high values; Guizhou falls into the third quadrant (low-low), and with a significant local Moran's I, forms a clustering of low values; Sichuan falls into the fourth quadrant (high-low), presenting a polarised pattern where high values are surrounded by low values; Provinces such as Anhui, Jiangxi, Hubei, Hunan, Chongqing and Yunnan are mainly distributed near the third quadrant, but their local Moran's I is not significant, and no clear clustering pattern has formed.

3.4 Analysis of Influencing Factors Based on a Principal Component Regression Model

To identify the key influencing factors of the green development efficiency (TE) of the cultural industry in the Yangtze River Economic Belt, this study employs a principal component regression method. First, a multicollinearity diagnosis is conducted on the candidate explanatory variables; then, principal components are extracted to construct a composite index; finally, a regression analysis is performed to examine the relationship between the composite index and efficiency, with robustness tests conducted using univariate regression.

3.4.1 Diagnosis of Multicollinearity

The variance inflation factors (VIFs) for the five explanatory variables are shown in Table 6. The VIFs for both the urbanisation rate and per capita GDP are well over 10, indicating the presence of severe multicollinearity.

Table 6 Variance Inflation Factors (VIF) of Explanatory Variables

Variable	VIF	1/VIF
urban	13.74	0.073
rgdp	12.40	0.081
cult_share	3.80	0.263
energy_int	1.81	0.552
cult_gap	1.78	0.562
Mean VIF	6.70	–

3.4.2 Principal Component Analysis

Multicollinearity exists between the urbanisation rate (urban), the share of value added in cult_share and energy_int. Principal component analysis was performed on these three variables. The KMO statistic was 0.68 and the Bartlett's sphericity test yielded a p-value of<0.01, indicating the presence of common factors among the variables and confirming the suitability of principal component analysis.

Table 7 Eigenvalues and Variance Explained

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	2.305	1.846	0.768	0.768
Comp2	0.459	0.224	0.153	0.922
Comp3	0.235	.	0.078	1.000

Table 8 Principal component load matrix

Variable	Comp1	Comp2	Comp3
urban	0.587	0.477	-0.654
cult_share	0.601	0.285	0.747
energy_int	-0.543	0.831	0.119

In accordance with the principle that eigenvalues should be greater than 1, the first principal component (PC1) was selected, explaining 76.8% of the variance. The load matrix (Table 8) shows that PC1 is positively correlated with urban and cult_

share, and negatively correlated with energy_int.

3.4.3 Principal Component Regression Results

Ordinary least squares regression was performed with PC1 as the explanatory variable and technical efficiency(TE)as the dependent variable, the results are shown in Table 9.

Table 9 Principal Component Regression Results

te	Coefficient	Standard error	t	P>t	[95%confidence	interval]
pc1	0.251	0.056	4.470	0.002	0.124	0.379
_cons	0.548	0.081	6.730	0.000	0.364	0.732

The regression coefficient of PC1 comes out as 0.251, and it is statistically significant at the 1% level (t = 4.47, p = 0.002). This means PC1 has a clear positive relationship with technical efficiency: a one-unit rise in PC1 corresponds to a rise of about 0.251 in technical efficiency. The model yields an R^2 of 0.690 and an adjusted R^2 of 0.656, and the F-test gives a p-value of 0.0015, suggesting that the model fits the data quite well overall.

3.4.4 Robustness Tests

To ensure the reliability of the conclusions, univariate regressions were conducted separately for urban, cult_share and energy_int (Table 10). Additionally, the univariate results for per capita GDP and the urban-rural cultural consumption gap are presented for comparison.

Table 10 Results of Univariate Regression

Variable	Coefficient	Std.err.	p	R^2
rgdp	1.308	0.309	0.002	0.666
urban	1.139	0.331	0.007	0.569
energy_int	-0.985	0.431	0.048	0.368
cult_share	1.683	1.023	0.161	0.290
cult_gap	-0.299	0.447	0.521	0.047

The results show that the univariate regression findings are consistent with those of the principal component regression, thereby validating the robustness of the conclusions.

4. Research Conclusions and Policy Recommendations

4.1 Research Conclusions

The overall comprehensive efficiency of green development in the cultural industry along the Yangtze River Economic Belt remains relatively low, exhibiting a pattern of divergence characterised by “high efficiency in the lower reaches, weak technical capabilities in the middle reaches, and limited scale in the upper reaches”. In 2023, the average comprehensive technical efficiency of green development in the cultural industry across the 11 provinces and municipalities of the Yangtze River Economic Belt stood at 0.548, with only Shanghai, Jiangsu and Zhejiang in the lower reaches achieving positive values. In the middle reaches of the Yangtze River, the pure technical efficiency of the four provinces was below 0.8, whilst in the upper reaches, the average scale efficiency of the four provinces was merely 0.277. All non-efficient provinces and cities exhibited positive slack in environmental pollution, and the three provinces in the lower reaches of the Yangtze River showed redundant capital investment.

Spatial disparities stem primarily from the inter-group gap between the Lower Yangtze and Upper Yangtze regions, with the most pronounced differentiation occurring within the Upper Yangtze region itself. The overall Gini coefficient stands at 0.436, with inter-group contributions accounting for 65.95%—the primary source of overall variation. The inter-group Gini coefficient is highest between the Lower Yangtze and Upper Yangtze regions (0.612) and lowest between the Middle Yangtze and Upper Yangtze regions (0.355). The Gini coefficient within the Upper Yangtze region is the highest (0.456), far exceeding that within the Lower Yangtze region (0.231) and the Middle Yangtze region (0.114), with a stark contrast in efficiency

between Sichuan (0.603) and Guizhou (0.021).

The pattern of spatial agglomeration is polarized, mainly showing “high–high” and “low–low” clusters with clear local polarization. The global Moran’s I is 0.527 ($p = 0.001$), confirming a strong positive spatial autocorrelation. Shanghai, Jiangsu, and Zhejiang belong to the high-high agglomeration zone. Guizhou falls into the low-low agglomeration zone, with a local Moran’s I of 0.519 ($p = 0.042$). Sichuan, however, presents a high-low polarization (local Moran’s $I = -0.093$, $p = 0.022$), meaning that a high-value region is surrounded by low-value areas, and no spillover effect is observed.

Urbanisation rates and industrial scale exert a positive driving force,whilst energy intensity imposes a negative constraint. Principal component regression indicates that PC1, derived from urbanisation rates, the share of cultural industry value added (positively correlated) and energy intensity (negatively correlated), has an efficiency coefficient of 0.251 ($p = 0.002$). Univariate tests show that for every 1-unit increase in per capita GDP, efficiency rises by 1.308, whilst for every 1-unit increase in energy intensity, efficiency falls by 0.985.

4.2 Policy Recommendations

4.2.1 Policy Coordination

Strengthen regional coordination and implement targeted measures to address the structural imbalance characterised by “high efficiency in the lower reaches, weak technology in the middle reaches, and limited scale in the upper reaches”, and develop digital cultural industries in a manner tailored to local conditions. We should build upon the high-quality development strategy for the Yangtze River Economic Belt during the 15th Five-Year Plan period, driving the green development of the cultural industry from “localised leadership” towards “overall coordination”. Firstly, we should develop the digital cultural industry in a manner suited to local conditions. Drawing on the differing cultural resource endowments of the middle and lower reaches of the Yangtze River, the lower reaches should focus on developing the metaverse, AI-generated animation, micro-dramas, online literature IPs, and digital films and television series; the middle reaches should focus on developing digital books, audiobooks, and cultural data assets; the lower reaches of the Yangtze River should prioritise the development of animation, gaming, digital music and ecological imaging, thereby reducing physical resource consumption at source. Secondly, a Green Collaborative Development Fund for the Cultural Industry should be established to provide key support for enhancing pure technical efficiency in the middle reaches of the Yangtze River and expanding effective scale in the upper reaches, with priority investment directed towards digital infrastructure such as cloud rendering and copyright blockchain. Thirdly, a “Green Development Efficiency” leader scheme should be implemented, offering green finance and land use incentives to Shanghai, Jiangsu and Zhejiang, whilst requiring them to establish low-carbon standards for the digital cultural sector. Fourthly, implement a special campaign to enhance scale efficiency in the upper reaches of the Yangtze River, addressing the challenge of insufficient scale through the cross-provincial co-construction of digital cultural industrial parks and the sharing of computing power and copyright resources.

4.2.2 Business Model Transformation

Establish a green cultural consumption and production system to overcome the negative constraints of energy intensity and the spatial polarisation pattern. In response to the negative impact—where a one-unit increase in energy intensity leads to a 0.985-unit decrease in efficiency—as well as the polarised clustering patterns of “high-high” and “low-low” regions, as highlighted in, the green transformation of the entire cultural industry chain must be accelerated. Firstly, vigorously develop new digital cultural business models, replacing high-energy-consuming traditional physical formats with animation, online literature, micro-dramas, digital books and audiobooks to reduce energy consumption per unit of added value. Secondly, promote low-carbon cultural venues by implementing energy consumption quota management for theatres, museums and similar facilities, whilst encouraging the use of renewable energy and green building materials. Thirdly, foster eco-digital convergence models—such as virtual performances in the metaverse and cultural and creative products linked to carbon credits—to internalise environmental externalities as economic benefits. Fourthly, establish green cultural consumption demonstration zones in regions experiencing “high–low” polarisation, such as Sichuan, where consumption subsidies and green credits can be used to stimulate demand for green digital cultural products and break the deadlock in spillover effects.

4.2.3 Technology Empowerment

Break through the bottleneck of pure technical efficiency in the middle reaches of the Yangtze River, and drive green technologies from input redundancy towards output conversion. The average pure technical efficiency in the middle reaches of the Yangtze River is below 0.8, and the existence of capital input redundancy in the lower reaches indicates insufficient technological conversion capacity. Firstly, apply artificial intelligence (AI) technology to intelligent editing of micro-dramas and short videos, optimisation of animation rendering, and assisted translation of online literature, thereby reducing computing power consumption and idle labour; deploy metaverse technology to create virtual production, cloud exhibitions and digital cultural tourism spaces, replacing physical set construction with virtual scenes to reduce waste of building materials; introduce blockchain technology to establish a platform for the authentication and trading of cultural data assets, revitalising idle copyright resources, achieving cross-regional resource sharing and avoiding duplicate investment. Secondly, establish a pilot platform for green technologies in the cultural industry, focusing on the research and development of low-energy-consumption stage equipment, biodegradable cultural and creative materials, and intelligent temperature-controlled exhibition systems, to address the issue of technological investment yielding no tangible output. Thirdly, promote the integration of digitalisation and intelligent technologies with green technologies, utilising the Internet of Things (IoT) to monitor energy consumption in cultural enterprises in real time, and employing AI to optimise production scheduling and logistics routes, thereby reducing unintended by-products (such as exhibition waste and high-carbon exhibition set-ups). Fourthly, establish a green technology efficiency assessment mechanism, linking internal R&D expenditure to pollution reduction and energy consumption cuts, thereby preventing technological investments from going to waste.

4.2.4 Talent Support

Cultivate professionals with expertise in both green culture and technology to resolve the inherent contradiction where technology is effective but lacks scalability. In the Upper Yangtze region, pure technical efficiency stands at 1.0, whilst scale efficiency is merely 0.277, reflecting a talent structure skewed towards the technical side and lacking in industrial operational capabilities. Firstly, establish a micro-specialisation in green operations for digital culture, piloted in universities in Sichuan, Chongqing, Yunnan and Guizhou, to cultivate multi-skilled professionals proficient in the mass production of micro-dramas, audiobook recording pipelines, the operation of online literature platforms, carbon accounting and green supply chains. Secondly, implement the Green Culture Leadership Programme, dispatching operational talent familiar with digital cultural business models from high-efficiency regions in the lower reaches of the Yangtze River to take up temporary posts in the upper reaches, thereby facilitating the integration of technical expertise with local scale expansion. Thirdly, “green efficiency performance” will be incorporated into the cultural industry talent evaluation system. Those demonstrating significant achievements in energy conservation, consumption reduction and the development of green digital products will be given priority in professional title promotions and project funding, thereby stimulating talent vitality.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Central Committee of the Communist Party of China, & State Council. (2024, August 11). Opinions of the Central Committee of the Communist Party of China and the State Council on accelerating the comprehensive green transformation of economic and social development. Retrieved April 18, 2026, from https://www.gov.cn/zhengce/202408/content_6967663.htm

[2] Ministry of Ecology and Environment, Ministry of Culture and Tourism, China Federation of Literary and Art Circles, & Chinese Writers’ Association. (2025, May 30). Guidelines on further strengthening ecological culture development. Retrieved April 18, 2026, from https://www.gov.cn/zhengce/zhengceku/202506/content_7026365.htm

[3] Central People’s Government of the People’s Republic of China. (2026, March 13). Outline of the 15th Five-Year Plan for national economic and social development of the People’s Republic of China. Retrieved April 18, 2026, from https://

www.gov.cn/yaowen/liebiao/202603/content_7062633.htm

[4] Shi, X. L., Cheng, Y., Zhang, J. N., et al. (2026). A study on the evolution of cultural industry agglomeration and its effects on green economic transformation in the Yellow River Basin: Based on micro-enterprise data. World Geographical Research, 1–17. Advance online publication.

[5] Chen, L., Chen, Z. M., & Yang, H. Y. (2025). The value origins, internal logic and practical strategies of new-quality productive forces empowering the high-quality development of the cultural industry. Journal of Shenyang University of Technology (Social Sciences Edition), 18(3), 332–339.

[6] Luo, L. (2025). The internal logic, practical dilemmas and implementation pathways of new-quality productive forces empowering the green development of the cultural industry. Xinjiang Social Sciences, (1), 142–150.

[7] Huang, X., & Hu, A. G. (2025). Dimensions of understanding green productivity, China's innovation and practical prospects: With a discussion on how new-quality productivity itself constitutes green productivity. Journal of Beijing University of Technology (Social Sciences Edition), 25(1), 56–67.

[8] Zheng, B. Y., & Chen, Y. P. (2024). Cultural new-quality productive forces: Scientific connotation, structural elements and development pathways. Journal of North Minzu University, (6), 141–149.

[9] Wan, Q. (2025). Evaluation of cultural industry efficiency and research on influencing factors in the Yangtze River Economic Belt. National Circulation Economy, (8), 145–148.

[10] Li, J., Gao, R. Z., Wang, X. Q., et al. (2025). Spatio-temporal variation and impact mechanisms of green development efficiency in resource-based cities along the Yangtze River Economic Belt. Resources and Environment of the Yangtze River Basin, 34(6), 1193–1207.

[11] Chen, L., Hu, L. J., Shi, J. W., et al. (2025). A study on the impact of regional economic policies on green technological innovation performance: A quasi-natural experiment based on the development strategy of the Yangtze River Economic Belt. Theory and Practice of Systems Engineering, 45(9), 2831–2852.

[12] Jiang, Z. R., Chen, Y. J., Pang, J. P., et al. (2025). The impact of the Yangtze River Economic Belt development strategy on ecological efficiency along the belt. Acta Ecologica Sinica, 45(12), 5839–5852.

[13] Zhang, X. L., Xuan, Z. Y., & Yi, J. B. (2025). Implementation outcomes of the Yangtze River Economic Belt development strategy and key measures for the 15th Five-Year Plan. Reform, (8), 64–76.

[14] Yuan, L., Zhang, S. Q., He, W. J., et al. (2025). A study on the impact of digitalisation and intelligence levels in the Yangtze River Economic Belt on green total factor productivity. Resources and Environment of the Yangtze River Basin, 34(5), 923–936.

[15] Shan, B. Y., Li, W. J., Wang, L. N., et al. (2025). The impact of the cultural industry on regional green development and spatial spillover effects: A case study of Shandong. Science, Technology and Economy, 38(6), 96–100.

[16] Hu, Y. L. (2025). New-quality productive forces empowering the high-quality development of the digital cultural industry: Internal logic, value implications and practical pathways. Comparative Studies on Cultural Innovation, 9(8), 76–80.

[17] Li, J. G. (2024). Practical pathways for the high-quality development of the cultural industry guided by the new development philosophy. Comparative Studies on Cultural Innovation, 8(28), 115–119.

[18] Qiu, G. Q. (2024). The concept of green development leading a new trend in the cultural industry: A review of A Reader on Green Cultural Development. Science, Technology and Publishing, (11), 139.

[19] Zhou, J. X., & Zhu, X. P. (2025). Annual academic report on research into China's cultural industries 2024. Journal of Shenzhen University (Humanities and Social Sciences Edition), 42(1), 46–58.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Supply Chain Concentration and Corporate Default Risk: Evidence from China

Renyuan Gu, Yangyue Huang, Qiying Li, Zihong Zhong, Qisong Wang*

School of Finance and Trade, Zhuhai College of Science and Technology, Zhuhai, 519041, China

Corresponding author: Qisong Wang, wangqisongmust@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Supply chain concentration constitutes a pivotal governance structure shaping corporate operational resilience. Using data on Chinese A-share listed firms from 2001 to 2022, this study estimates expected default probability via the KMV model and employs a random-effects Tobit model to correct for left-censoring in the dependent variable. Panel estimation results reveal that elevated supply chain concentration significantly reduces corporate default probability, with this effect operating through both customer concentration and supplier concentration dimensions. To address endogeneity, we construct an instrumental variable using the number of Ming Dynasty post stations and apply the control function approach to correct for bias. The corrected estimates confirm the risk mitigation effect of supply chain concentration.

Keywords: Supply Chain Concentration; Customer Concentration; Default Risk; Supplier Concentration

Published: Jun 3, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1476

1. Introduction

Against the backdrop of global value chain restructuring and the elevation of industrial chain security to national strategic priority, supply chain stability and resilience have become critical variables influencing macroeconomic performance. In recent years, from global supply chain disruptions triggered by the COVID-19 pandemic (Li et al., 2023), to raw material shortages caused by geopolitical conflicts, and bottlenecks in critical components, frequent supply chain risk events have not only impacted micro-level corporate operations but also evolved into threats to systemic financial stability through credit linkages and risk contagion mechanisms among enterprises (Ağca et al., 2022). Notably, government support policies play an important role in alleviating corporate financing constraints and enhancing survival capabilities, yet such supportive effects may exhibit nonlinear boundaries (Han et al., 2024). As node enterprises within supply chain networks, corporate defaults reflect not merely individual financial distress but may trigger chain reactions through debt linkages, amplifying risks across entire industries and even the financial system (Xie et al., 2023). Therefore, identifying the intrinsic relationship between supply chain structural characteristics and corporate default risk holds significant practical importance for preventing and resolving major financial risks and safeguarding industrial and supply chain security.

However, existing literature has primarily focused on internal governance dimensions such as equity structure and board characteristics, while research on how informal institutional networks (e.g., chambers of commerce and industry associations as social capital vehicles) influence cross-border resource allocation has grown increasingly prevalent (Wang et al., 2025). Nevertheless, examination of supply chains as external governance mechanisms remains insufficient. Supply chain

relationships serve as carriers of implicit contracts between enterprises and possess governance functions; deep collaboration can reduce information asymmetry and enhance operational stability, yet excessive dependence may amplify operational uncertainty and risk contagion effects. Under this divergence, how does supply chain structure influence corporate default risk? Answering this question not only helps clarify the microeconomic consequences of supply chain management but also provides empirical evidence and policy references for improving corporate governance mechanisms and strengthening risk prevention systems.

Existing research exhibits three limitations regarding core conclusions, methodology, and identification strategies. First, in terms of core conclusions, Fang et al. (2023) found that supply chain concentration is significantly positively correlated with default risk, meaning higher concentration leads to higher expected default probability, with mechanisms operating through crowding-out effects on innovation investment and performance. Other studies have focused on the impact of customer concentration on cash flow risk (Mihov & Naranjo, 2017), stock price crash risk (Ma et al., 2020), corporate risk-taking (Cao et al., 2021), and loan contract terms (Campello & Gao, 2017), yet direct examination of the relationship between supply chain concentration and default risk remains insufficient. Although Shen and Chen (2023) studied default risk among supply chain enterprises, they focused on the U-shaped effect of short-term debt for long-term use, treating customer effects merely as moderating variables without directly testing the net effect of supply chain concentration itself on default risk. Second, methodologically, existing literature generally employs OLS or standard linear panel regression (e.g., Fang et al., 2023), ignoring the substantial left-censoring of default probability at zero, which leads to inconsistent estimators. Third, regarding identification strategy, supply chain concentration is likely endogenous; firms with lower default risk may self-select into more concentrated supply chain networks, yet existing research has failed to effectively address this endogeneity bias or introduce appropriate instrumental variables for causal identification.

Addressing these gaps, the marginal contributions of this study are threefold. First, regarding core conclusions, this study finds that increased supply chain concentration significantly reduces corporate default probability, supporting the risk mitigation effect of concentration and providing new empirical evidence for the debate on supply chain structure. Second, methodologically, this study systematically applies the RE-Tobit model to this field for the first time, effectively correcting estimation bias caused by left-censoring of default probability; simultaneously, it adopts the control function approach proposed by Blundell and Powell (2004), constructing an instrumental variable using the historical geographic variable of Ming Dynasty post station numbers, obtaining more reliable causal effects through two-stage estimation. Third, regarding mechanism analysis, this study decomposes the effect into customer concentration and supplier concentration, revealing complementary transmission channels through downstream demand stability and upstream supply security in mitigating default risk, filling the gap in micro-level mechanism evidence.

2. Theories and Hypothesis

Social capital theory and resource dependence theory posit that firms are embedded within networks composed of upstream suppliers and downstream customers (Portes, 1998). Within these networks, trust and repeated interactions serve as implicit contracts, restraining opportunistic behavior and reducing monitoring costs. For firms operating in volatile environments, concentrated supply chains can function as stable resource channels: core suppliers ensure production continuity through prioritized capacity allocation, while key customers provide predictable cash flows, thereby smoothing revenue volatility. From this perspective, supply chain concentration can reduce transaction costs, compress information asymmetry, and ultimately increase default distance while lowering default probability (PD).

However, resource dependence theory also cautions that excessive reliance on a limited number of partners creates lock-in effects. Once dominant suppliers or customers encounter liquidity shocks, distress will transmit along the supply chain, amplifying the focal firm's default risk (Guo et al., 2024). Furthermore, network homophily may trap firms in information echo chambers, distorting resource allocation and amplifying vulnerability (McPherson et al., 2001). Through this lens, higher concentration narrows default distance and increases default probability.

Given the theoretical ambiguity regarding the net effect of supply chain concentration on default risk, this study proposes the following competing hypotheses,

H1a: Ceteris paribus, increased supply chain concentration enlarges default distance, thereby reducing corporate default risk.

H1b: Ceteris paribus, increased supply chain concentration narrows default distance, thereby exacerbating corporate default risk.

As a composite indicator, supply chain concentration masks differentiated transmission mechanisms between upstream and downstream. Drawing on existing literature regarding decomposition of supply chain network structure (Patatoukas, 2012; Wu and Yao, 2023), this study disaggregates SCCD into customer concentration (CC) and supplier concentration (SC) to separately test their risk mitigation pathways.Downstream channel: Demand stability mechanism. Customer concentration reflects the concentration degree of a firm's revenue streams. When firms depend on a limited number of core customers, although they face demand interruption risks, deep collaborative relationships can reduce sales volatility, shorten accounts receivable cycles, and provide stable cash flow expectations (Campello & Gao, 2017). This demand stability directly improves firms' short-term solvency and interest coverage ratios, thereby expanding default distance. Accordingly, this study proposes,

H2a: Increased customer concentration reduces corporate default probability by enhancing demand stability.

Upstream channel: Supply security mechanism. Supplier concentration reflects the concentration degree of a firm's input sources. Establishing long-term relationships with core suppliers ensures prioritized raw material supply, reduces search costs and bargaining costs, and provides capacity guarantees during supply chain disruptions. This supply security reduces production stagnation risks and inventory accumulation costs, stabilizing firms' cost structures and profit margins, thereby reducing default probability. Accordingly, this study proposes,

H2b: Increased supplier concentration reduces corporate default probability by enhancing supply security.

These two channels are not mutually exclusive but complementary: downstream demand stability safeguards cash inflows, while upstream supply security controls cost expenditures, both jointly affecting corporate default distance. If H1a holds, both H2a and H2b should be simultaneously validated; if H1b holds, at least one channel should be insignificant or exhibit opposite direction.

Supply chain concentration is likely endogenous: firms with lower default risk may self-select into more concentrated supply chain networks, or omitted governance quality factors may simultaneously influence both. To obtain causal effects, instrumental variables satisfying both relevance and exogeneity conditions must be identified.

This study employs the number of Ming Dynasty post stations in prefecture-level cities where firms are located as the instrumental variable, with the following logical foundations,

Relevance condition. Ming Dynasty post stations served as core nodes for information transmission and material allocation during the imperial period, with site selection following principles of geographic continuity and accessibility to all directions, prioritizing transportation hubs with active commerce and trade. These historical transportation nodes shaped the logistics network foundations and commercial agglomeration traditions of contemporary cities, making firms in these regions more likely to form close supply chain cooperative relationships. Therefore, the number of Ming Dynasty post stations is positively correlated with contemporary supply chain concentration.

Exogeneity condition. The layout of Ming Dynasty post stations was determined by military defense and administrative transmission requirements, bearing no direct relationship with contemporary corporate default risk. Post station numbers remain constant over time, absorbing permanent unobservable city characteristics such as geographic endowments and historical hub status, which may influence contemporary economic structures but do not directly affect individual corporate default behavior. Moreover, as a historical variable, post stations are uncorrelated with endogenous disturbances in contemporary supply chain policy, satisfying the exclusion restriction.

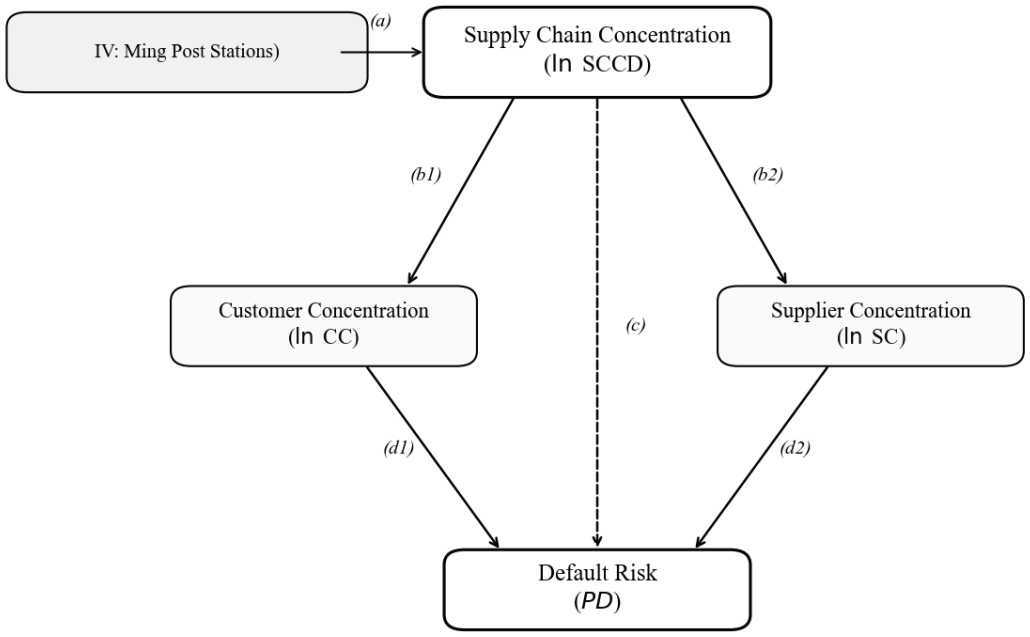
Based on the above analysis, this study proposes,

H3: The number of Ming Dynasty post stations is significantly positively correlated with supply chain concentration, satisfying the relevance requirement for instrumental variables; and this instrumental variable is exogenous to corporate default risk, satisfying the exclusion restriction.

Figure 1 illustrates the mechanism transmission pathways through which supply chain concentration affects default risk.

This figure presents the conceptual framework and empirical strategy of this study. Pathway (a) represents the first-stage regression, using the number of Ming Dynasty post stations as the instrumental variable (IV) to identify the exogenous variation in supply chain concentration (ln SCCD). Pathway (c) depicts the direct effect of supply chain concentration on corporate default probability (PD) (H1). Pathways (b1) and (b2) decompose supply chain concentration into customer concentration (lnCC) and supplier concentration (lnSC). Pathways (d1) and (d2) represent two mechanism transmission channels: customer concentration affects default risk through demand stability (H2a), while supplier concentration operates through supply security (H2b). All regressions control for net profit margin (lnNPM), cash ratio (lnCR), equity-to-debt ratio (lnE/D), and year fixed effects.

Figure 1 Mechanism Transmission of Supply Chain Concentration on Default Risk



Controls: Profitability (ln NPM), Cash Ratio (ln CR), Equity-to-Debt (ln E/D), Year FE

3. Empirical Strategy

3.1 Sample Selection and Data Sources

This study examines Chinese A-share listed companies from 2001 to 2022, constructing an unbalanced panel dataset. Sample selection follows these criteria: (1) financial sector listed companies are excluded to avoid interference from their distinctive capital structures; (2) ST and ST firms are excluded; (3) observations with missing key variables are excluded; (4) to mitigate extreme value effects, continuous variables are trimmed at the 1st and 99th percentiles, meaning observations outside this range are directly deleted. After these screenings, the final sample comprises 31,561 firm-year observations. Data sources include the CSMAR and Wind databases.

3.2 Variable Declaration

The dependent variable in this study is corporate default risk. Following Bharath and Shumway (2008) and Deren X and Jinsong L (2022), we select PD_{it} as the indicator for measuring corporate default risk. The specific calculation formulas are as follows.

First, corporate default probability PD_{it} is obtained by transforming default distance DD_{it} through the standard normal cumulative density function:

$$PD_{it}=N(-DD_{it}) \tag{1}$$

In equation (1), N(\cdot) denotes the standard normal cumulative density function.

Second, default distance DD_{it} measures the degree to which corporate asset market value deviates from the default point, calculated as:

$$DD_{it}=\frac{V_{it}-VD_{it}}{V_{it}\cdot\sigma_{V,it}} \tag{2}$$

In equation (2), V_{it} represents corporate asset market value, obtained through simultaneous iterative solution of equations (4) and (5);VD_{it} represents the default point (F_{it}) Calculated as the aggregate of current liabilities and 0.5 times the non - current liabilities at the end of the year.; and $\sigma_{V,it}$ epresents corporate asset value volatility.

Furthermore, corporate asset value volatility $\sigma_{V,it}$ is calculated through weighted combination of equity value volatility$\sigma_{E,it}$ and debt volatility $\sigma_{D,it}$.

$$\sigma_{V,it}=\frac{VE_{it}}{VE_{it}+VD_{it}}\cdot\sigma_{E,it}+\frac{VD_{it}}{VE_{it}+VD_{it}}\cdot\sigma_{D,it}=\frac{VE_{it}}{VE_{it}+VD_{it}}\cdot\sigma_{Eit}+\frac{VD_{it}}{VE_{it}+VD_{it}}\cdot\left(0.05+0.25\cdot\sigma_{E,it}\right) \tag{3}$$

In equation (3), VE_{it} represents the equity market capitalization, which is defined as the product of the total number of outstanding shares and the stock price at year end ; $\sigma_{E,it}$ represents equity value volatility, measured by the monthly stock return standard deviation over the previous year; $\sigma_{D,it}$ represents debt volatility, approximated based on equity volatility as $\sigma_{D,it}=0.05+0.25\left(0.05+0.25\cdot\sigma_{E,it}\right)$.

Finally, solving for V_{it} and $\sigma_{V,it}$ requires simultaneous solution of the following two equations. Equation (4) is the Black-Scholes option pricing equation:

$$VE_{it}=V_{it}\cdot N\left(\frac{\ln\left(\frac{V_{it}}{F_{it}}\right)+\left(r_{it}+0.5\sigma_{V,it}^{2}\right)\cdot T}{\sigma_{V,it}\cdot\sqrt{T}}\right)-F_{it}\cdot e^{-r_{it}\cdot T}\cdot N\left(\frac{\ln\left(\frac{V_{it}}{F_{it}}\right)+\left(r_{it}+0.5\sigma_{V,it}^{2}\right)\cdot T}{\sigma_{V,it}\cdot\sqrt{T}}-\sigma_{V,it}\cdot\sqrt{T}\right) \tag{4}$$

In equation (4), VE_{it} represents equity market value; r_{it} represents the risk-free rate, calculated using one-year time deposit interest rates weighted by time; and T denotes debt maturity, set at 1 year.

Equation (5) is the volatility relation.

$$\sigma_{E,it}=N\left(\frac{\ln\left(\frac{V_{it}}{F_{it}}\right)+\left(r_{it}+0.5\sigma_{V,it}^{2}\right)\cdot T}{\sigma_{V,it}\cdot\sqrt{T}}\right)\cdot\frac{V_{it}}{E_{it}}\cdot\sigma_{V,it} \tag{5}$$

Through simultaneous solution of equations (4) and (5) using iterative methods, corporate asset market value V_{it} and asset value volatility $\sigma_{V,it}$ are obtained. Substituting these into equation (2) yields default distance DD_{it} and subsequently default probability PD_{it} through equation (1). A larger DD_{it} value indicates higher safety margin between corporate asset value and the default point, meaning lower default probability PD_{it} and lower default risk.

The explanatory variable is supply chain concentration (SCCD). Following Qiang W and Yu-xiu Y (2023), this study measures enterprise supply chain concentration using the mean of the sum of supplier concentration (SC) and customer concentration (CC), with specific calculation formulas as follows.

$$SC=\frac{\sum_{i=1}^{n}SC}{n} \tag{5}$$

$$CC=\frac{\sum_{i=1}^{n}CC}{n} \tag{6}$$

$$SCCD=\frac{SC+CC}{2} \tag{7}$$

In equations (5) and (6), supplier concentration and customer concentration are expressed as averaged sums. Supplier concentration is measured by the proportion of total annual procurement from the top five suppliers relative to total annual procurement; customer concentration is measured by the proportion of total annual sales to the top five customers relative to total annual sales; and equation (7) calculates listed company supply chain concentration as the mean of customer concentration and supplier concentration.

This study introduces net profit margin (lnNPM), cash ratio (lnCR), and equity to debt ratio (lnE/D) as control variables. Additionally, year dummy variables are included to control for macro-level unobservable shocks. See Tables 1 and 2 for details.

Table 1 Descriptive statistics of variables

variable	Obs	Mean	Std. dev.	Min	Max
PD	31,561	0.202	0.282	0	1
lnSCCD	31,561	3.385	0.516	1.723	4.456
lnCC	31,561	3.202	0.795	0.231	4.584
lnSC	31,561	3.390	0.579	1.635	4.556
lnNPM	31,561	0.190	0.168	-4.605	0.504
lnCR	31,561	-0.935	1.163	-4.692	2.297
LnE/D	31,561	0.416	0.968	-2.847	2.999
IV	31,561	3.582	3.838	0	19
Year	31,561	2016.279	5.297	2001	2022

Table 2 Description of the main variables

variable	Variable Description
PD	Default Probability Ratio
lnSCCD	The mean of the sum of the procurement proportions from the top five suppliers and the sales proportions from the top five customers (truncated logarithm
lnCC	The proportion of sales to the top five customers in total annual sales (truncated logarithm)
lnSC	The proportion of procurement value from the top five suppliers to the total annual procurement value (truncated logarithm)
lnNPM	The ratio of Net profit to operating receipt (logarithm truncated)
lnCR	The ratio of the sum of monetary funds and trading financial assets to current liabilities (truncated logarithm)
LnE/D	The ratio of the Total Owner's Equity to liabilities (truncated logarithm)
IV	The number of Ming Dynasty post stations in the prefecture-level cities where the enterprises are located.
Year	Year dummy variable

3.3 Model Specification

Because corporate default probability PD_{it} is left-censored (Left-censored), meaning some observations are compressed to zero, ordinary panel regression would yield biased estimates. Therefore, this study adopts the random-effects Tobit (RE-Tobit) model as the baseline estimation method. This model introduces random individual effects c_i to capture time-invariant firm heterogeneity and assumes that c_i is independent of the explanatory variables. The model specification is as follows.

$$PD_{it}^{*} = \beta_0 + \beta_1 lnSCCD_{it} + \gamma Controls_{it} + c_i + \mu_{it} \quad (1)$$

$$PD_{it} = \max(0, PD_{it}^{*}) \quad (2)$$

$$c_i | X_i \sim N(0, \sigma_c^2), \ \mu_{it} | X_i, c_i \sim N(0, \sigma_\mu^2) \quad (3)$$

where PD_{it}^{*} is the latent variable of corporate default probability; PD_{it} is the observed KMV default probability; $Controls_{it}$ is the vector of control variables; c_i is the random individual effect, $c_i \sim N(0, \sigma_c^2)$, which is independent of the explanatory variables; μ_{it} is the random disturbance term, $\mu_{it} \sim N(0, \sigma_\mu^2)$; σ_c^2 and σ_μ^2 are the variances of the individual effect and the random error, respectively.

4. Empirical Results

4.1 Baseline Regression Analysis

Table 3 reports the baseline regression results. Columns (1) and (2) present the pooled Tobit estimation results, while columns (3) and (4) present the random-effects Tobit (RE-Tobit) estimation results. The core explanatory variable lnSCCD exerts a significantly negative effect on the dependent variable in columns (1) and (3), and this result remains significant after control variables are included in columns (2) and (4). The RE-Tobit results in columns (3) and (4) indicate the presence of significant individual effects, rendering the RE-Tobit model superior to the pooled Tobit model. As a reference, the pooled Tobit results in columns (1) and (2) likewise support the reliability of the core conclusions.

Table 3 Baseline Regression Result

Model DV	Tobit PD (1)	Tobit PD (2)	RE-Tobit PD (3)	RE-Tobit PD (4)
lnSCCD	-0.0612*** (-18.22)	- 0.0224*** (-8.18)	-0.0577*** (-9.69)	-0.0351*** (6.36)
lnNPM		-0.0148 (-1.62)		-0.0600*** (-5.62)
lnCR		0.0213*** (13 .19)		0.0109*** (4.86)
lnE/D		-0.1799*** (-76.71)		-0.1262*** (-30.57)
_cons	0.2953*** (19.85)	0. 2388*** (8.51)	0.1908*** (7.56)	0.2422*** (0.60)
VCE	robust	robust	Bootsrap	Bootsrap
Year FE	YES	YES	YES	YES
Id FE	NO	NO	NO	NO
Obvs	31561	31561	31561	31561
Left-censord	5	5	5	5
Right-censord	334	334	334	334
Log-likelihood	-4154.5221	2060.3482	4240.3146	6574.2564
wald	χ^2=2376.44 p=0.0000	χ^2=13411.75 p=0.0000	χ^2=2732.12 p=0.0000	χ^2=3707.80 p=0.0000
rho			0.5598	0.4194

Note: *, **, and *** indicate significance at the 10%,5%, and 1% levels, respectively; the values in parentheses are t/z values.

4.2 Mechanism Inspection Analysis

To reveal the micro-level transmission channels through which supply chain concentration affects default risk, this study decomposes supply chain concentration into customer concentration (LnCC) and supplier concentration (lnSC). Table 4 reports the RE-Tobit estimation results. Columns (1) and (2) show that customer concentration exerts a significantly negative effect on default risk; columns (3) and (4) show that supplier concentration likewise significantly reduces default risk. Both dimensions contribute to the risk mitigation effect, indicating that downstream demand stability and upstream supply security are complementary transmission mechanisms. As shown in the mechanism test results in Table 4, to reveal the transmission pathways through which supply chain digitalization affects default risk, this study employs the RE-Tobit model for mechanism testing, focusing on the mediating roles of supply chain concentration and supply chain stability.

Table 4 Mechanism Testing Results

Model DM	RE-Tobit PD (1)	RE-Tobit PD (2)	RE-Tobit PD (3)	RE-Tobit PD (4)
lnCC	-0.0356*** (-8.17)	-0.0226*** (-6.06)		
lnSC			-0.0398*** (-7.51)	-0.0238*** (-5.30)
_cons	0.1027*** (5.23)	0.1920*** (11.59)	0.1279*** (5.49)	0.2033*** (9.92)
Controls	NO	YES	NO	YES
Id Year	YES	YES	YES	YES
Id FE	NO	NO	NO	NO
VCE	Bootsrap	Bootsrap	Bootsrap	Bootsrap
Obvs	31561	31561	31561	31561
Left-censord	5	5	5	5
Right-censord	334	334	334	334
Log-likelihood	4212.9576	6567.8599	4196.5981	6551.9578
Wald	χ^2 = 2390.59 p=0.0000	χ^2= 3795.94 p= 0.0049	χ^2 = 2819.61 p=0.0000	χ^2 = 3766.95 p=0.0000
Rho	0.5602	0.4179	0.5603	0.4190

Note: *, **, and *** indicate significance at the 10%,5%, and 1% levels, respectively; the values in parentheses are t/z values.

4.3 Robustness Test

Table 5 reports the robustness test results. To ensure that our conclusions are not influenced by model specification or variable measurement, we conduct tests along two dimensions. First, we change the estimation method by replacing the RE-Tobit model with OLS. Second, we replace the dependent variable by substituting default probability PD with its logarithmic form lnPD. Columns (1) and (2) use PD as the dependent variable, while columns (3) and (4) use lnPD as the dependent variable. The core explanatory variable lnSCCD remains significantly negative across all columns, with significance at the 1% level in all cases. These results indicate that, regardless of whether linear estimation or logarithmic transformation is adopted, the risk-reducing effect of supply chain concentration on default risk remains stable, further corroborating the reliability of our baseline conclusions.

Table 5 Robustness test results

Model DM	OLS PD (1)	OLS PD (2)	OLS lnPD (3)	OLS lnPD (4)
lnSCCD	-0.0507*** (-6.63)	-0.0384*** (-5.60)	-0.2102*** (-6.93)	-0.15673*** (-5.86)
_cons	0.1584*** (5.26)	0.2182*** (8.05)	-2.7454*** (-21.21)	-2.4765*** (-21,71)
Controls	NO	YES	NO	YES
Year FE	YES	YES	YES	YES
Id FE	YES	YES	YES	YES
VCE	Robust	Robust	Robust	Robust
Obvs	31561	31561	31561	31561
R^2	0.1285	0.2036	0.1789	0.2391

Note: *, **, and *** indicate significance at the 10%,5%, and 1% levels, respectively; the values in parentheses are t/z values.

4.4 Endogenous Processing

Table 6 reports the endogeneity-corrected estimation results. To alleviate potential endogeneity bias arising from omitted variables and reverse causality, this study adopts the control function (CF) approach proposed by Blundell and Powell (2004). In the first stage, using the historical geographic variable, the number of Ming Dynasty post stations in the prefecture-level city where the firm is located, as the instrumental variable (IV), we regress the endogenous variable lnSCCD via OLS and extract the residual term denoted as Cf1. In the second stage, the residual term is incorporated into the RE-Tobit model for estimation.

First-stage results show that the instrumental variable exerts a significantly negative effect on lnSCCD, and the first-stage F-statistic is 29.14, far exceeding the Stock-Yogo weak instrument test critical value of 16.38, thereby rejecting the weak instrument hypothesis and satisfying the relevance requirement of the instrumental variable. Second-stage RE-Tobit-CF estimation shows that the coefficient on the core explanatory variable lnSCCD is significantly negative at the 5% level, and the first-stage residual term Cf1 is significantly positive, indicating that endogeneity indeed exists in the original baseline model. However, after correcting for the bias, the core conclusion that supply chain concentration reduces corporate default risk remains valid.

Table 6 Endogenousity test results

DV **Model** **Stage**	**(1)** **lnSCCD** **OLS** **First stage**	**(2)** **PD** **RE-Tobit-CF** **Seconed stage**
IV	-0.0028* (-2.77)	
lnSCCD		-0.7633* (-2.53)
Cf1		0.7283* (2.41)
_cons	3.4599*** (113.57)	2.7628* (2.64)
Controls	YES	YES
Year FE	YES	YES
Id FE	NO	NO
VCE	Robust	Bootstrap
Obvs	31561	31561
Left-censord		5
Right-censord		334
Log-likelihood		6577.1656
R^2/Wald	R^2=0.0226	χ^2 =9929.54
First-stage F	29 .14***	

Note: *, **, and *** indicate significance at the 10%,5%, and 1% levels, respectively; the values in parentheses are t/z values.

5. Conclusions

Based on data from Chinese A-share listed enterprises spanning 2001 to 2022, this study systematically examines the impact of supply chain concentration on listed company default risk. RE-Tobit estimation results indicate that elevated supply chain concentration significantly reduces corporate default probability, and this conclusion remains valid after OLS robustness tests and control function endogeneity correction. Mechanism analysis further reveals that both customer concentration and supplier concentration contribute to the risk mitigation effect, indicating that downstream demand stability and upstream supply security jointly guide firms in avoiding financial distress.

Based on these findings, this study proposes the following policy recommendations. For enterprises, they should prudently evaluate supply chain structure and establish long-term strategic cooperative relationships with core suppliers and major customers to balance efficiency and resilience. For regulators, they should strengthen supply chain information disclosure requirements for listed companies, particularly concentration data regarding top five customers and suppliers, enabling markets to timely identify potential supply chain risk contagion nodes. For financial institutions, they may incorporate supply chain concentration and its changing trends into corporate credit rating systems as auxiliary indicators for default risk early warning.

Funding

This research was generously supported by the Guangdong Province Philosophyand Social Science Foundation (Grant No. GD24XLJ06).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Ağca, S., Babich, V., Birge, J. R., & Wu, J. (2022). Credit shock propagation along supply chains: Evidence from the CDS market. Management Science, 68(9), 6506-6538. https://doi.org/10.1287/mnsc.2021.4174

[2] Li, Z., Pei, S., Li, T., & Wang, Y. (2023). Risk spillover network in the supply chain system during the COVID-19 crisis: Evidence from China. Economic Modelling, 126, 106403. https://doi.org/10.1016/j.econmod.2023.106403

[3] Xie, X., Shi, X., Gu, J., et al. (2023). Examining the contagion effect of credit risk in a supply chain under trade credit and bank loan offering. Omega, 115, 102751. https://doi.org/10.1016/j.omega.2022.102751

[4] Han, Y., Han, L., Liu, C., & Wang, Q. (2024). How does government R&D subsidies affect enterprises' viability? An investigation on inverted U-shaped relationship[wangqisongmust@163.comwangqisongmust@163.comJ]. Finance Research Letters, 106235. https://doi.org/10.1016/j.frl.2024.106235

[5] Han, Y., Wang, Q., Xiang, X., & Liu, M. (2026). Executives' Green Experience, Green Innovation and Corporate Environmental Performance: A Machine Learning-Based News Sentiment Analysis in China. International Review of Economics & Finance, 104956. https://doi.org/10.1016/j.iref.2026.104956

[6] Wang, Q., Tu, X., Liu, C., & Han, Y*. (2025). How informal institutions influence entrepreneurs' cross-border location and industry choices: Empirical evidence from Macao's associations. Finance Research Letters, 108527. https://doi.org/10.1016/j.frl.2025.108527

[7] Xie, D., & Liu, J. (2022). Free cash flow productivity and default risk: Evidence from A-share listed companies. Journal of Financial Research, 510(12), 168–186. http://www.jryj.org.cn/EN/abstract/abstract1132.shtml

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

[9] Wang, C., Li, T., Zhu, L., et al. (2026). Conductor of supply chain orchestra: Chain intermediary shareholders and enterprise co-innovation. Economic Analysis and Policy, 91, 1203–1225. https://doi.org/10.1016/J.EAP.2026.04.015

[10] Qiao, R., & Zhao, L. (2026). Passive information blockade and proactive information sharing: Impact of asymmetric digitalization on supply chain financing risks and role of governance. Supply Chain Management: An International Journal, 31(3), 424–445. https://doi.org/10.1108/SCM-11-2025-1128

[11] Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. The Review of Financial Studies, 21(3), 1339–1369. https://doi.org/10.1093/rfs/hhn044

[12] Jia, F., Zhang, T., Chen, L., et al. (2026). How do supply chain finance platforms adopt blockchain to achieve better service performance? An information processing perspective. Supply Chain Management: An International Journal, 31(3), 446–462. https://doi.org/10.1108/SCM-07-2025-0710

[13] Shen, H., & Chen, J. (2023). Research on U-shaped relationship between short-term debt for long-term use and supply chain enterprise default risk: Evidence from Chinese listed firms. PloS one , 18(10), e0293284. https://doi.org/10.1371/JOURNAL.PONE.0293284

[14] Anonymous. (2010). Oliver Wyman report: Supply chain risk: Companies must develop credit scoring capabilities to predict supplier defaults. Manufacturing Close-Up. https://www.enterrasolutions.com/insights/supply-chain-risks-whos-in-charge-and-what-are-they-looking-at

[15] Phonthanukitithaworn, C., & Ketkaew, C. (2026). Managing green supply chain performance through digital twin technology – insights from SMEs and large firms. Journal of Open Innovation: Technology, Market, and Complexity, 12(2), 100762. https://doi.org/10.1016/J.JOITMC.2026.100762

[16] Xue, J., & Qu, Y. (2026). How does digital supply chain finance affect the fluctuation of corporate cash flow? Based on panel data analysis of industry heterogeneity. International Review of Economics and Finance , 108, 105230. https://doi.org/10.1016/J.IREF.2026.105230

[17] Wang, Y., Liu, Y., Xie, Y., et al. (2026). Chain-master firms' digitalization and supply chain resilience: Evidence from China. Economic Modelling, 161, 107623. https://doi.org/10.1016/J.ECONMOD.2026.107623

[18] Dadsena, K. K., Rathore, P., Mathiyazhagan, K., et al. (2026). Adoption of circular economy practices in supply chains to enhance firm performance: Industrial symbiosis and ecological modernization theory perspective. The International Journal of Logistics Management, 37(3), 755–785. https://doi.org/10.1108/IJLM-06-2024-0341

[19] Tran, M. T. D., Thai, V. V., Duc, H. T. T., et al. (2026). The nexus between organisational culture and differentiation-based competitive advantage in garment firms in Vietnam: Does supply chain collaboration matter? The International Journal of Logistics Management, 37(3), 644–667. https://doi.org/10.1108/IJLM-03-2025-0169

[20] Omidian, A., Omidian, A., Mohajer, S., et al. (2026). Sustainable supply chain management, information technology, and firm performance: The mediating power of agility and supply chain integration. Tehnički glasnik, 20(2), 226–232. https://doi.org/10.31803/TG-20250315172439

[21] Rafiq, M., Yang, J., Bashar, S. (2026). Operational efficiency as a contingent, not universal mechanism: Evidence from Pakistani SME supply chains. Supply Chain Management: An International Journal, 31(3), 407–423. https://doi.org/10.1108/SCM-08-2025-0834

[22] Blundell, R., & Powell, J. L. (2004). Endogeneity in semiparametric binary response models. The Review of Economic Studies, 71(3), 655–679. https://doi.org/10.1111/j.1467-937X.2004.00299.x

[23] Campello, M., & Gao, J. (2017). Customer concentration and loan contract terms. Journal of Financial Economics, 123(1), 108–136. https://doi.org/10.1016/j.jfineco.2016.03.010

[24] Cao, Y., Dhaliwal, D. S., Li, Z., et al. (2021). Customer concentration and corporate risk-taking. Journal of Financial Stability , 100890. https://doi.org/10.1016/j.jfs.2021.100890

[25] Dhaliwal, D. S., Judd, J. S., Serfling, M., & Shaikh, S. (2016). Customer concentration risk and the cost of equity capital. Journal of Accounting and Economics, 61(1), 23–48. https://doi.org/10.1016/j.jacceco.2015.03.005

[26] Fang, G. Z., Meng, F. B., & Zuo, S. (2023). Research on the impact of supply chain concentration on corporate default risk. Industrial Technology & Economy, 42(9), 201–211. https://doi.org/10.3969/j.issn.1004-910X.2023.09.011

[27] Guo, C., Zhang, S., & Chen, S. (2024). Does carbon risk travel along the supply chain? Evidence from corporate default risk. Economics Letters, 236, 111602. https://doi.org/10.1016/j.econlet.2024.111602

[28] Ma, S., Ruan, X., Shen, C., & Wu, H. (2020). Customer concentration and stock price crash risk. Journal of Corporate Finance , 64, 101692. https://doi.org/10.1016/j.jcorpfin.2020.101692

[29] McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27, 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

[30] Mihov, M. T., & Naranjo, A. (2017). Customer-base concentration and cash-flow risk. Journal of Empirical Finance, 42, 1–23. https://doi.org/10.1016/j.jempfin.2016.11.006

[31] 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

[32] Portes, A. (1998). Social capital: Its origins and applications in modern sociology. Annual Review of Sociology, 24, 1–24. https://doi.org/10.1146/annurev.soc.24.1.1

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Substation Indoor Environment Monitoring System based on STM32

Lei Mei*, Sixiao Zhang, Rui Mei, Yulong Shi, Ji Ma

State Grid Jibei Electric Power Co., Ltd., Zhangjiakou Power Supply Company, Zhangjiakou, Hebei, 075000, China

Corresponding author: Lei Mei, 22126251@bjtu.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Aiming at the problems of low efficiency of manual monitoring of indoor environment in substation, untimely abnormal response and lack of intelligent control of equipment operation environment, an indoor environment monitoring system of substation based on STM32 single chip microcomputer is designed. The system takes STM32F103C8T6 as the main control core, integrates temperature and humidity, smoke, harmful gas, water leakage detection and other types of sensors, realizes real-time acquisition of environmental parameters, local display, over-limit alarm and remote data transmission functions, and adapts to the strong electromagnetic working environment of substation through anti-interference design. The test results show that the system has stable operation, high data acquisition accuracy and timely alarm response, which can effectively ensure the safe and stable operation of substation indoor equipment and improve the intelligent level of substation environmental management.

Keywords: STM32; Substation; Environmental Monitoring; Sensors

Published: Jun 8, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1435

1. Introduction

As the core hub of power systems, substations' indoor equipment operating environments directly impact grid reliability. Issues such as abnormal temperature/humidity levels, harmful gas leaks, smoke/fire incidents, and water accumulation/leaks can lead to insulation degradation in high-voltage equipment, short circuits, corrosion, or even safety accidents. Traditional manual inspections and scheduled monitoring methods suffer from data delays, missed inspections, and delayed emergency responses, failing to meet smart grid requirements for automated and intelligent substation operations. With rapid advancements in embedded and IoT technologies, microcontroller-based environmental monitoring systems (Li Renpeng, Fu Chuannan & Huang Miao, 2025) have been widely adopted in industrial applications. The STM32 series microcontrollers offer robust processing capabilities, abundant peripheral resources, low power consumption, and excellent interference resistance, making them ideal for complex substation environments. This paper presents an STM32-based indoor environmental monitoring system that enables round-the-clock real-time monitoring of multiple parameters, automatic anomaly alerts, and remote supervision, providing technical support for safe substation operations.

2. System Overall Design

2.1 System Design Requirements

To meet the practical requirements for substation indoor environmental management, the system must fulfill the following

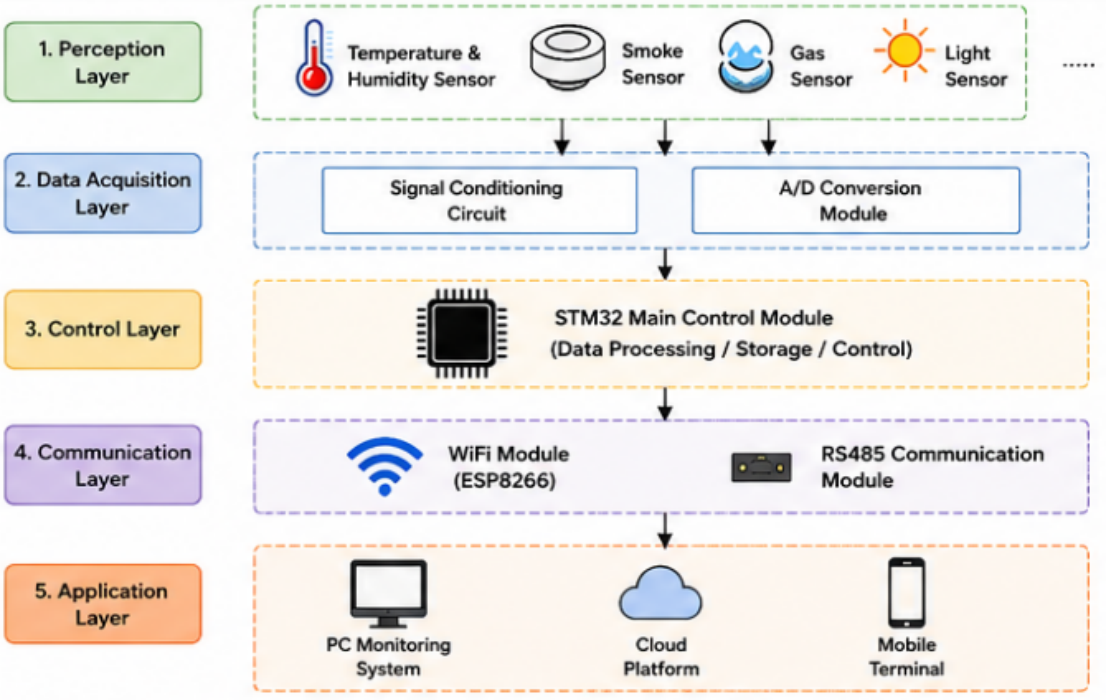
functions:

real-time acquisition of critical environmental parameters such as indoor temperature and humidity, smoke concentration, harmful gases (carbon monoxide, hydrogen sulfide), and water leakage status (Zhang Bo, 2025); local real-time display of monitoring data to visually present the environmental conditions; immediate activation of audible and visual alarms when environmental parameters exceed safety thresholds to alert maintenance personnel; remote wireless transmission of monitoring data supporting real-time viewing and storage on backend computers; and robust electromagnetic compatibility to withstand strong electromagnetic interference in substations, ensuring stable system operation.

2.2 System Architecture

The system adopts the hierarchical architecture design of ‘ front-end acquisition-main control processing-output control-remote transmission ‘. The whole system is divided into six parts: main control module, sensor acquisition module, display module, alarm module, wireless communication module and power module.

Figure 1 The overall structure diagram of the substation indoor environment monitoring system based on STM32



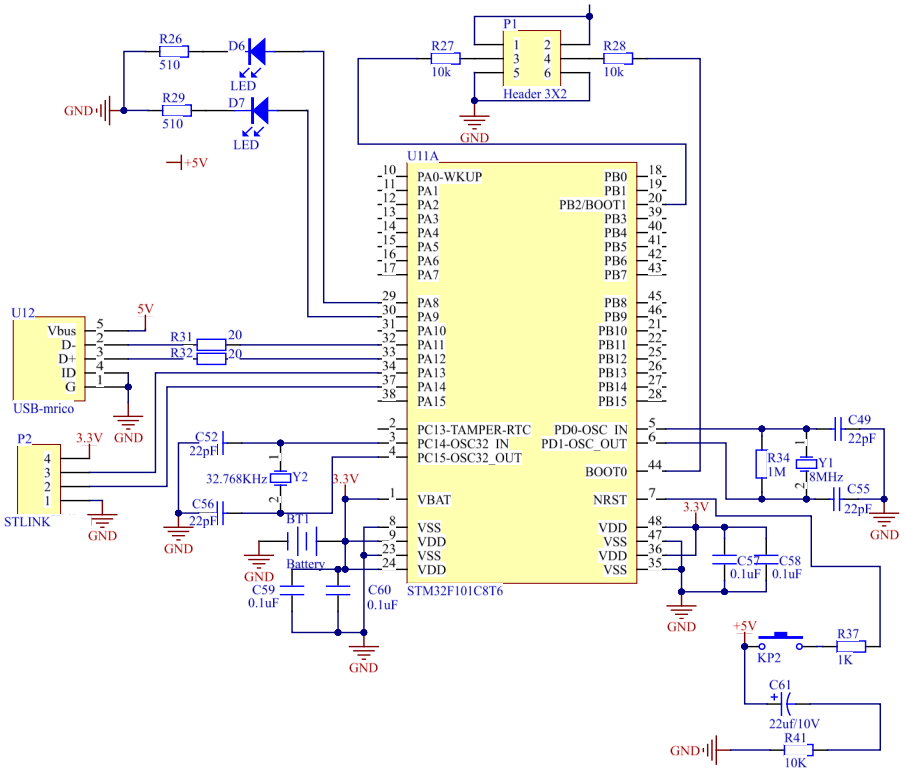
The working process of the system : the sensor acquisition module collects all kinds of environmental parameters in the substation room in real time, and transmits the analog signal / digital signal to the STM32 main control module ; after filtering, analog-digital conversion and data analysis of the signal, the main control unit drives the display module to complete the local display of the data on the one hand, and on the other hand, the monitoring data is compared with the preset threshold. If the data exceeds the limit, the alarm module is started immediately. At the same time, the processed data is uploaded to the remote monitoring platform through the wireless communication module to realize remote real-time monitoring.

3. System Hardware Design

3.1 Master Module

The main control unit employs the STM32F103C8T6 microcontroller, which is based on the ARM Cortex-M3 core (Xu Junjie, Wei Jinghe, Liu Guozhu, He Jian & Zhang Zheng, 2025). With a maximum clock frequency of 72 MHz, it features 128 KB of Flash memory and 20 KB of SRAM, along with comprehensive GPIO interfaces, an ADC analog-to-digital conversion interface, and a serial communication interface, enabling efficient data acquisition, processing, and communication for multiple sensors. Operating within a voltage range of 2.0–3.6 V, the chip offers low power consumption, high stability, and support for various low-power modes, making it ideal for long-term uninterrupted operation in substations.

Figure 2 STM32 minimum system

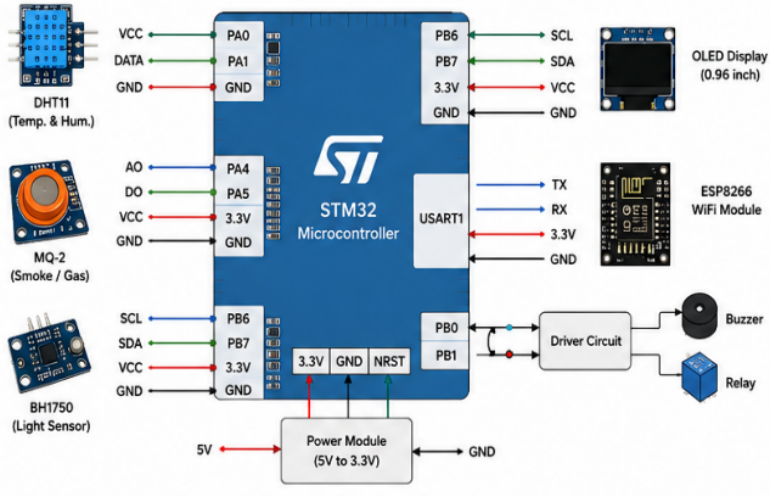


3.2 Sensor Acquisition Module

According to the environmental monitoring requirements of the substation, four types of high-precision sensors are selected to achieve multi-parameter comprehensive acquisition:

Temperature and humidity sensor: DHT11 digital temperature and humidity sensor is selected, and single bus communication protocol is adopted. It can collect 0-50 °C temperature and 20 % -90 % RH humidity data. The accuracy meets the requirements of indoor environment monitoring in substation, and the interface is simple and anti-interference is strong. Smoke sensor: MQ-2 smoke sensor is selected. Based on semiconductor gas sensor, it can detect smoke and combustible gas concentration in real time. It has fast response speed and high sensitivity, and is suitable for early warning of substation fire. Harmful gas sensor: MQ-135 gas sensor is selected to detect toxic and harmful gases such as carbon monoxide and hydrogen sulfide, so as to avoid safety risks caused by equipment failure and gas leakage. Leakage sensor: The contact type leakage sensor is used to detect the water accumulation on the ground and the leakage state of the pipeline in the substation, and the switch signal is output to prevent the water from damaging the power equipment.

Figure 3 System hardware connection schematic diagram



In the design of sensor circuit, signal filtering and photoelectric isolation circuit are added to weaken the influence of strong electromagnetic interference on the acquisition signal and ensure the accuracy of data acquisition.

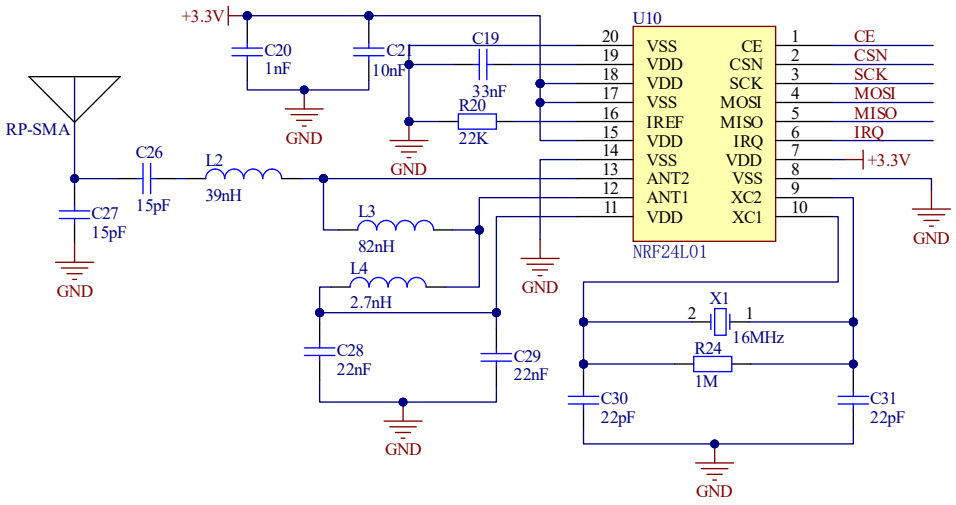
3.3 Display and Alarm Module

Display module: The 0.96-inch OLED display screen is connected to the STM32 through the I2 C interface. It has the advantages of low power consumption, clear display, fast response, and large visual angle. It can display real-time monitoring data such as temperature and humidity, gas concentration, and water leakage status, which is convenient for on-site operation and maintenance personnel to view. The alarm module is composed of buzzer and LED warning light, which is connected with the main control through GPIO interface. When the monitoring data exceeds the preset safety threshold, STM32 outputs the control signal to drive the buzzer to shout and the LED lamp to flicker, so as to realize the sound and light double alarm and improve the abnormal reminding effect.

3.4 Wireless Communication Module

The NRF24L01 module (Guo Jingbo, Zhang Chao, Zhao Weixing, Zhang Haidong & Zhou Qingxiang, 2018) was selected for remote data transmission. This module supports serial port transparent transmission mode, enabling rapid LAN integration to upload STM32-processed environmental monitoring data to the remote monitoring platform while receiving threshold adjustment and device control commands from the platform, achieving bidirectional communication. With its compact size, stable transmission performance, and low cost, the module meets the wireless data transmission requirements for substation interiors.

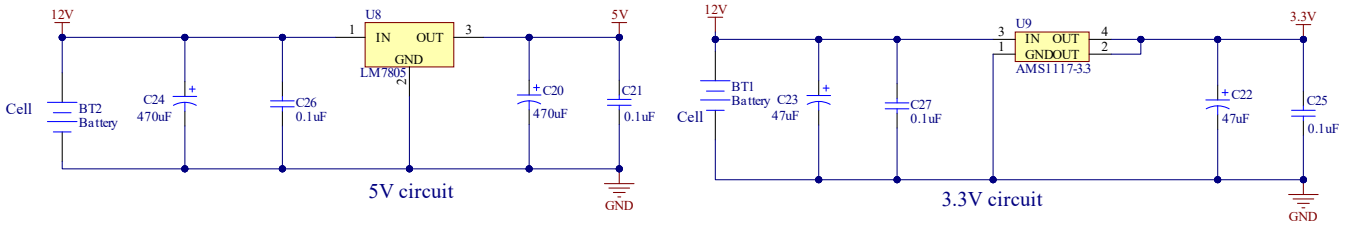
Figure 4 Wireless communication schematic diagram



3.5 Power Module

It is powered by 12 V DC power supply, which is converted to 5 V and 3.3 V stable voltage by step-down circuit, which is powered by sensor, WiFi module and STM32 main control unit respectively. The voltage regulator chip, filter capacitor and anti-reverse protection circuit are added to the circuit to improve the stability of the power supply, avoid the influence of voltage fluctuation on the operation of the system, and adapt to the on-site power supply environment of the substation.

Figure 5 Power supply circuit



3.6 Electromagnetic Compatibility Design

In view of the strong electromagnetic interference environment of the substation, multiple anti-interference measures are

adopted in the hardware design : the main control circuit and the sensor acquisition circuit are separately wired to reduce signal crosstalk ; magnetic beads and capacitive filtering are added to the communication line ; the key circuit is protected by shielding cover ; the power interface is connected to the transient voltage suppression diode to prevent the surge voltage from damaging the device and ensure the stable operation of the system in the complex electromagnetic environment.

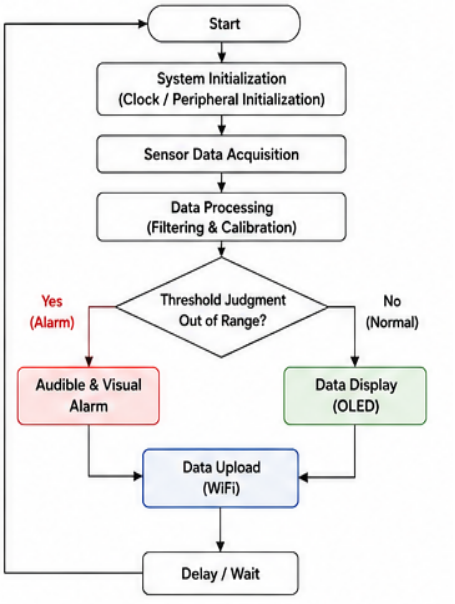
4. System Software Design

The system software is based on the Keil MDK development environment, using C language programming, following the modular design idea, to achieve data acquisition, processing, display, alarm, communication and other functions. The main program flow is: system power-on initialization → sensor data acquisition → data filtering processing → local display → threshold comparison judgment → normal display / overrun alarm → data remote transmission → loop execution.

4.1 Main Program Design

After the system is powered on, the initial configuration of STM32 system clock, GPIO interface, ADC, serial port, OLED display and WiFi module is completed. Then enter the main cycle, regularly start the sensor data acquisition, read the sensor output signal through the interrupt mode, perform mean filtering and error correction processing on the collected original data, and improve the data accuracy; the processed data is pushed to the OLED display screen, and compared with the preset temperature, humidity and gas concentration threshold. If the threshold is exceeded, the alarm will be triggered. Finally, the data is packaged and uploaded to the remote platform through the WiFi module, and the cycle is executed after the completion of a round of monitoring process.

Figure 6 System software flow chart



4.2 Data Acquisition Subroutine

For the communication protocols used by different sensors, it is necessary to write corresponding exclusive data acquisition subroutines to ensure the accuracy and real-time performance of data reading. Specifically : DHT11 temperature and humidity sensor based on single bus communication protocol, the subroutine through strict timing control (such as reset signal and response detection, accurate delay with high and low level switching) to complete the data bit by bit read, finally resolve the temperature and humidity value ; for gas sensors such as MQ-2 and MQ-135, the output of the two is positively correlated with the concentration of the analog voltage signal. The subroutine needs to configure the built-in ADC module of STM32 (including channel selection, sampling time setting and resolution calibration). After converting the analog quantity into digital quantity, the corresponding gas concentration value is calculated according to the calibration curve or fitting formula. The leakage sensor uses digital level signal to judge the leakage state, and the subroutine continuously polls or reads the level of the specified IO port in real time through external interruption. Once the effective level jump is detected (such as high and low indicating leakage), the corresponding processing logic (such as setting alarm sign, closing water valve or sending alarm

message) is triggered immediately, so as to realize the rapid response and disposal of abnormal leakage.

4.3 Alarm and Communication Subroutine

The alarm subroutine adopts the interrupt trigger mode, which aims to realize the rapid response to abnormal situations and maximize the safety and reliability of the monitoring system. When any monitoring data (such as temperature and humidity, gas concentration or water leakage state) exceeds the preset safety threshold, the corresponding sensor acquisition subroutine will immediately trigger a hardware or software interrupt, thereby interrupting the current main program. Performing routine tasks (such as data storage, interface refresh, etc.), preferentially jump to the alarm service function. In this function, the system first controls the active buzzer through the GPIO port to continuously output high-frequency sounds, and at the same time drives the high-brightness LED lamp to flash at a specific frequency to form a significant sound and light warning signal to remind the on-site operation and maintenance personnel to dispose of it in time. If the exception persists, the alarm signal will remain active and will not be allowed to be overwritten by the main program or other low-priority tasks. Only when one of the following two exit conditions is satisfied, the alarm state will be cleared: First, all monitoring indicators are restored to the normal range for several consecutive sampling periods, and the system automatically determines that the fault is eliminated, and then the sound and light output is turned off and the main program cycle is restored. Secondly, the operation and maintenance personnel perform manual reset at the scene through buttons or remote instructions, forcing the termination of the current alarm and re-entering the monitoring state. The dual clearing mechanism not only supports automatic operation, but also retains the flexibility of manual intervention, effectively avoiding the long-term failure of the alarm due to signal jitter or residual error.

The communication subroutine is developed using the AT command set of the ESP8266 WiFi module (Zhou Yuanyuan, Sun Zhen & Wang Xuhong, 2026), serving as the core component responsible for reliably transmitting monitoring data to the remote platform and parsing and executing 下行 commands. The subroutine first sends a series of initialization commands to the ESP8266 via the serial port, including setting the operating mode to Station (or AP + Station mode), configuring the target access point's SSID and password, enabling multi-connection or single-connection mode, and establishing a TCP or MQTT connection with the remote server. Once the communication link is established, the subroutine packages the collected data—including temperature, humidity, gas concentration, and water leakage status—according to predefined communication protocols (e.g., JSON format, Modbus-TCP, or custom frame structures). These data frames are then transmitted to the ESP8266 via the serial port using either direct transmission mode or sequential command encapsulation, and forwarded to the cloud platform. To ensure transmission reliability, the subroutine incorporates mechanisms such as timeout retry, automatic reconnection upon connection loss, and data buffer retransmission, maintaining data integrity even under wired or network fluctuations. Meanwhile, the subroutine continuously monitors the ESP8266's serial port data through polling or interrupts. Upon receiving control commands from the platform (e.g., "trigger alarm," "reset sensor," or "adjust sampling frequency"), it immediately decodes the command code and activates the corresponding functional module to execute the required operation, enabling bidirectional closed-loop control between the monitoring node and the remote platform.

4.4 Filtering Algorithm Design

In order to eliminate the random error and instantaneous disturbance of the data caused by the complex electromagnetic environment, the noise of the sensor itself and the fluctuation of the ambient temperature and humidity in the industrial field, the system uses the sliding mean filtering algorithm to digitally filter the original collected data to improve the stability and reliability of the monitoring data. Specifically, the algorithm maintains a fixed-length FIFO data queue for each sensor channel (e.g., a queue length of 10 or 20, depending on the sampling frequency and interference characteristics). Each time, the newly collected data is pushed into the tail of the queue, and the old data of the head of the queue is moved out of the queue. Then, all the data samples in the current queue are preprocessed-firstly, the maximum and minimum values are sorted according to the value size, and the maximum and minimum values are eliminated (to eliminate the occasional pulse interference or spike noise), and then the remaining intermediate samples are arithmetically averaged. The calculation results are used as the effective data after this filtering for subsequent concentration conversion, over-limit judgment and upload processing. This sliding window filtering mechanism can not only effectively suppress high-frequency random interference

and periodic noise, but also track the output faster than the ordinary average filtering when the real value of the input signal changes step by step, avoiding excessive delay and signal lag. Compared with simple clipping filtering or median filtering, sliding mean filtering achieves a good balance between smoothness and real-time performance, which significantly improves the repeatability and long-term stability of the data collected by the system under complex working conditions, and provides more reliable raw data support for subsequent over-limit alarm, trend analysis and remote monitoring.

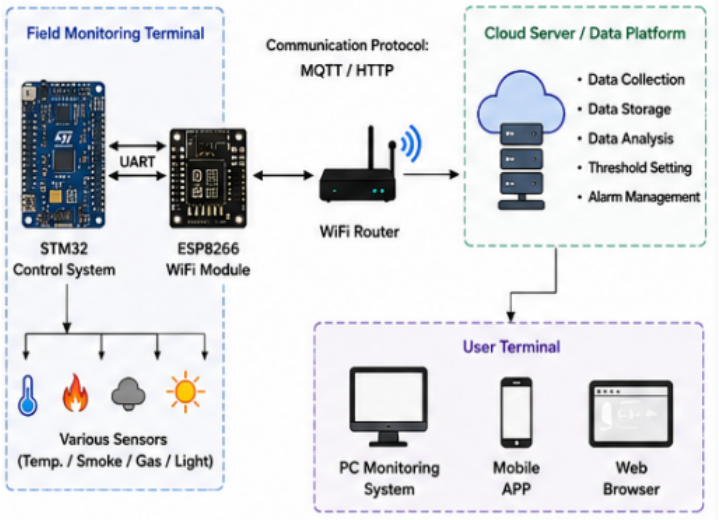
5. System Test and Result Analysis

In order to verify the performance of the system, the indoor simulation test environment of the substation is built, and the system function, data accuracy, stability and alarm response are comprehensively tested. The ambient temperature is 25 °C and the humidity is 45 % RH, and the conventional working state of the substation is simulated.

5.1 Functional Test

After the system is powered on and the initialization is completed, the OLED display immediately refreshes the current environmental monitoring data in real time with a clear font and a reasonable layout, including the temperature and humidity values collected by DHT11, the gas concentration estimation values output by MQ-2 and MQ-135, and the status information of the water leakage sensor. The display screen is stable without flicker, and the data update cycle is consistent with the acquisition frequency, which directly reflects the normal operation state of the system. In order to further verify the response ability of the system under different environmental conditions, a number of simulation tests were carried out in turn. By changing the temperature and humidity conditions of the local environment through the heater and the humidifier, the system can quickly identify the numerical changes after data acquisition and dynamically update on the OLED. The lighter is used to release a small amount of butane gas to simulate the smoke leakage scene. The gas concentration curve rises obviously, and the system immediately perceives the over-threshold trend. Drop a small amount of water into the water leakage sensor probe, and its output level jumps. At the moment of the above-mentioned abnormal triggering, the system can accurately determine the over-standard or water leakage state, and start the sound and light alarm-buzzer to issue a high-frequency warning sound under the drive of the alarm subroutine of the interrupt priority within 1 second. The LED warning light flashes synchronously, and the response is timely and the action is clear. At the same time, the WiFi communication module based on ESP8266 automatically accesses the preset wireless routing network after power on, and maintains a long connection state. There is no disconnection or authentication failure during the test. All filtered monitoring data are packaged in a custom JSON format, and are transmitted to the WiFi module through the serial port and stably pushed to the remote monitoring platform. By comparing the platform receiving records with the local data one by one, it is verified that the data transmission has no packet loss and no disorder, and the end-to-end delay is controlled at the millisecond level to meet the real-time monitoring requirements. In summary, the data acquisition, anomaly detection, alarm linkage and wireless communication of the system work stably and reliably, and all functional indicators meet the project design requirements.

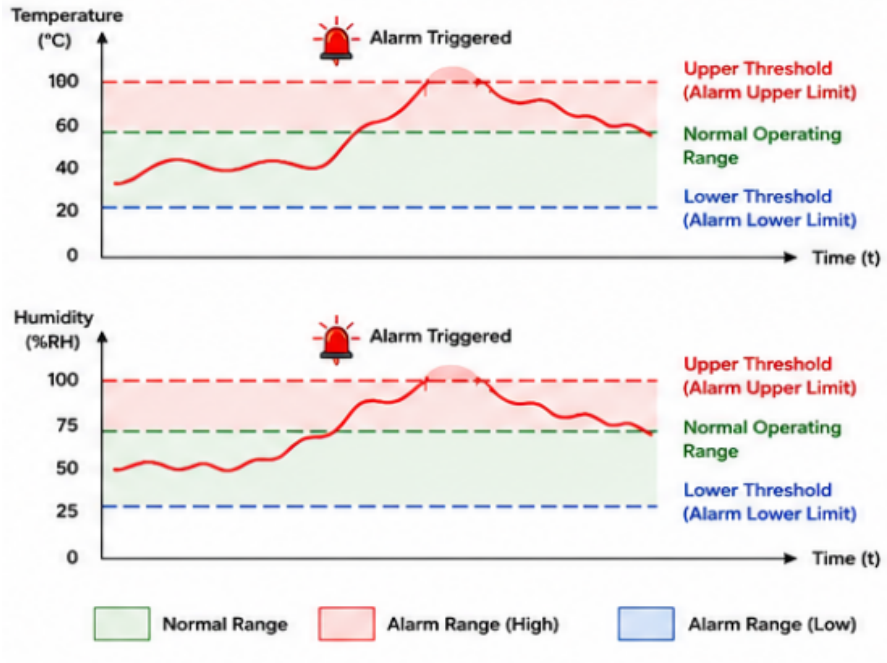
Figure 7 System communication and data transmission structure diagram



5.2 Data Accuracy Test

In order to ensure the accuracy and reliability of the monitoring data of the system, the standard temperature and humidity meter (accuracy of ± 0.1 °C, ± 1 % RH) and the standard gas detector are used as the reference equipment, and the data collected by the system are compared and tested synchronously under the same environmental conditions. During the test, several typical operating points (such as low temperature and low humidity, normal temperature and normal humidity, high temperature and high humidity, and different gas concentration scenarios) were selected for long-term continuous comparison records. The experimental results show that the maximum absolute error between the temperature data measured by the system and the standard temperature and humidity meter does not exceed ± 0.5 °C, the humidity error is controlled within ± 3 % RH, and the relative error of gas concentration detection does not exceed ± 5 %. The above error indexes are better than the allowable error requirements proposed in the relevant technical specifications of substation indoor environment monitoring (usually temperature error ≤ ± 1 °C, humidity error ≤ ± 5 % RH, gas concentration error ≤ ± 10 %). The achievement of this accuracy level benefits from the effective application of high-precision sensor selection, optimized signal conditioning circuit design and sliding mean filtering algorithm, which effectively suppresses the influence of electromagnetic interference and random noise on the sampling results. Therefore, the data of temperature, humidity and gas concentration collected and reported by the system have good accuracy, repeatability and long-term stability, which can be used as a reliable basis for substation environmental safety assessment, early warning of equipment operating conditions and remote operation and maintenance decision-making.

Figure 8 Environmental parameter monitoring and alarm mechanism diagram



5.3 Stability and Anti-Interference Test

Under the strong electromagnetic interference environment of the simulated substation, the system runs continuously for 72 hours without problems such as crash, data disorder, and communication interruption. The monitoring data fluctuation is small, the alarm function triggers accurately, and the electromagnetic compatibility design effect is remarkable. It can adapt to the complex industrial environment of the substation and has good long-term operation stability.

6. Conclusion and Foresight

The substation indoor environment monitoring system based on STM32 designed in this paper takes STM32F103C8T6 as the core, integrates multi-type sensors and wireless communication technology, and realizes the functions of all-weather real-time monitoring, local display, over-limit alarm and remote monitoring of substation indoor multi-environment parameters. The hardware structure of the system is simple, the cost is low, the anti-interference is strong, and the software design is modular and easy to maintain. The test results show that the system runs stably, the data is accurate, and the response is

timely. It can effectively solve the disadvantages of traditional substation environmental monitoring, improve the intelligent level of substation environmental management, and ensure the safe operation of power equipment. The follow-up research can further optimize the system function : add more environmental parameter monitoring modules and expand the monitoring dimension ; integrate into the Internet of Things cloud platform to realize data cloud storage, historical query, and abnormal report generation ; combined with mobile communication technology, the remote early warning push of mobile phone APP is realized, which further improves the convenience and practicability of the system and promotes the construction and development of intelligent substation.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Li Renpeng, Fu Chuannan & Huang Miao. (2025). Design of substation environment monitoring system based on Internet of Things. Mechanical Engineering & Automation, 54(05), 197-200.

[2] Zhang Bo. (2025). Automatic environmental monitoring method for 220kV substation infrastructure site under 5G network architecture. Microcomputer Applications, 41(07), 252-255.

[3] Xu Junjie, Wei Jinghe, Liu Guozhu, He Jian & Zhang Zheng. (2025). Design and Implementation of a PCIe RP System Based on the Cortex-M3 Core. Integrated Circuits and Embedded Systems, 25(08):74–80. https://doi.org/10.20193/j.ices2097-4191.2025.0051.

[4] Guo Jingbo, Zhang Chao, Zhao Weixing, Zhang Haidong & Zhou Qingxiang. (2018). Design of a fire pump inspection and monitoring system based on the nRF24L01 wireless module. Fire Science and Technology, 37(09),1228–1231.

[5] Zhou Yuanyuan, Sun Zhen & Wang Xuhong. (2026). Design of a Smart Home System Based on Cloud-Edge-Device Collaboration—With ESP32/ESP8266 as the Core. Internet of Things Technology, 16(09),65–69. https://doi.org/10.16667/j.issn.2095-1302.2026.09.015.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

The Internationalisation Challenges of Chinese SMEs: Formation Mechanisms and Governance Pathways

Hongyu Cao[1], Weixiang Gan[2]*

1. Zhenguigu Technology Co., Ltd. No. 55, Qingxiang Road, East of Jiuganqu, Guangyang Economic Development Zone, Langfang, Hebei, 065000, China

2. Lincoln University College, Petaling Jaya, Selangor Darul Ehsan-47301, Malaysia.

Corresponding author: Weixiang Gan, Gabriel1995@qq.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Against the backdrop of expanding foreign trade, a rising share of private enterprises and rapid growth in cross-border e-commerce, the internationalisation of Chinese SMEs shows a dual pattern: more market entry points and deeper capability constraints. Evidence from the General Administration of Customs, the Ministry of Industry and Information Technology, the World Bank, the WTO, UNCTAD, ADB and EU regulatory materials, combined with gradual internationalisation theory, global value-chain analysis and the born-global literature, indicates that the difficulties facing SMEs do not stem from a lack of overseas demand. They arise from a structural mismatch between firm capabilities and external rules. Low-price orders can lock firms into lower positions in global value chains; data-intensive compliance rules raise fixed costs; insufficient trade finance worsens cash-flow pressure; platform-based internationalisation weakens the accumulation of customer assets; and weak organisational capability prevents experience from becoming sustained competitiveness. Governance should therefore shift from entry indicators, such as export value, store numbers and trade-fair participation, to capability indicators, such as profit margins, collection cycles, compliance completeness, repurchase rates, financing coverage and after-sales response. Sectoral compliance tools, scenario-based trade finance, public services for industrial clusters and stronger internal governance can help Chinese SMEs move from 'products going global' to 'capabilities going global'.

Keywords: SMEs; Internationalisation; Global Value Chains; Trade Finance; Compliance Costs

Published: Jun 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1485

1. Introduction

The internationalisation of small and medium-sized enterprises (SMEs) is an important component of high-quality development in China's foreign trade, but it should not be reduced to export expansion or the growth of cross-border platform transactions. In 2025, China's total goods trade reached RMB 45.47 trillion, while private enterprises accounted for 57.3% of total imports and exports (General Administration of Customs, 2026). These figures show that private and SME-related business entities have become increasingly important in the foreign-trade system. At the same time, the World Bank's China Enterprise Survey reports that only 4% of firms in the China sample export directly, below the 8% observed in East Asia and the Pacific and the 9% observed in upper-middle-income economies (World Bank, 2025). This gap indicates that the cross-border movement of products does not automatically mean that firms directly control overseas customers, channels, pricing,

compliance or after-sales capabilities.

Theoretically, SME internationalisation involves at least three sets of capabilities. The first is market knowledge and commitment: the gradual internationalisation model argues that firms reduce uncertainty in foreign markets through accumulated experience (Johanson & Vahlne, 1977). The second is ownership advantage and internalisation capability: the eclectic paradigm suggests that firms without transferable technology, brands or organisational advantages find it difficult to obtain stable returns from internationalisation (Dunning, 1988). The third is value-chain position and governance capability: global value-chain research shows that buyer-driven structures can constrain suppliers' upgrading space (Gereffi et al., 2005). Accordingly, SME internationalisation should not be treated as a single act of exporting; it should be understood as a process through which firms develop customer, channel, financing, compliance, after-sales and brand capabilities in overseas markets.

Existing studies explain SME internationalisation through different but complementary routes. Gradual internationalisation emphasises experience accumulation and risk control, whereas the born-global literature shows that some firms internationalise early and rapidly through innovation, networks and entrepreneurial capabilities (Oviatt & McDougall, 1994; Knight & Cavusgil, 2004; Cavusgil & Knight, 2015). For Chinese SMEs, the central question is not simply whether internationalisation is fast or slow, but whether the process leaves behind capabilities that can be replicated. If a firm obtains short-term orders mainly through low prices, a single customer or platform traffic, the quality of its internationalisation remains limited.

Evidence from the General Administration of Customs, the Ministry of Industry and Information Technology, the World Bank, the WTO, UNCTAD, ADB and EU regulatory materials points to a coexistence of aggregate expansion and capability insufficiency in the internationalisation of Chinese SMEs. Behind foreign-trade resilience lies a clear micro-level gap: orders do not necessarily generate profits; platform entry does not necessarily create customer assets; and export scale does not necessarily imply upgrading in global value chains. Focusing only on export value or platform transaction value can obscure operating costs related to payment periods, compliance, financing and after-sales services.

SMEs are also highly heterogeneous and should not be treated as a uniform group. Public statistics can reveal the scale of foreign trade, market structure, cross-border e-commerce and firm-survey results, but they cannot fully capture each firm's profit margin, compliance expenditure or payment arrangements. A small number of specialised and sophisticated firms have strong advantages in niche markets, while a large number of ordinary SMEs still face low-price competition, weak financing access and insufficient compliance capacity. Recognising this stratification is consistent with the research tradition on SME heterogeneity (Lu & Beamish, 2001; OECD, 2009).

The tension between macro-level trade resilience and micro-level firm constraints is therefore central to understanding the internationalisation difficulties of Chinese SMEs. Total foreign-trade volume, the share of private enterprises and the growth of cross-border e-commerce show that market opportunities remain available. Yet direct export participation, access to bank loans, the trade-finance gap and the tightening schedule of green rules indicate that firms can benefit from internationalisation only when they possess specific capabilities. Low-price competition, compliance costs, trade finance, platform dependence and organisational capability are not separate issues; they interact within the same operating process.

Table 1. Number and Quality Tiers of Chinese SMEs

Indicator	Publicly reported basis	Value	Interpretation
Registered SMEs	End of 2024	More than 60 million	A very large population; support cannot be confined to a small group of leading firms.
Technology-based and innovative SMEs	Public reports in 2025	More than 600,000	A pool of technological potential that still needs channels and compliance capabilities.
Specialised and sophisticated SMEs	Public reports in 2025	More than 140,000	Stronger capacity in niche markets.
Specialised and sophisticated "little giant" firms	Public reports in 2026	About 17,600	The leading group is expanding, but it remains a minority.

Source: Compiled from public releases by the Ministry of Industry and Information Technology, Xinhua News Agency and China Government Network.

2. Data Evidence

China's overall foreign trade continues to show considerable resilience. Total goods trade increased from RMB 43.85 trillion in 2024 to RMB 45.47 trillion in 2025, while exports rose from RMB 25.45 trillion to RMB 26.99 trillion (China Government Network, 2025; General Administration of Customs, 2026). Imports and exports by private enterprises increased from RMB 24.33 trillion to RMB 26.04 trillion, raising their share of total trade from 55.5% to 57.3%. Private enterprises have thus become a major body of foreign-trade growth, but they also bear more of the pressure from external demand fluctuations, trade frictions, payment-period risks and compliance costs. Meanwhile, China has a large SME population with marked quality stratification. By the end of 2024, more than 60 million SMEs were registered. By 2025-2026, there were more than 600,000 technology-based and innovative SMEs, more than 140,000 specialised and sophisticated SMEs, and about 17,600 specialised and sophisticated 'little giant' firms (MIIT-related releases, 2026). The group with technological, professional and niche-market capabilities is expanding, but it remains small relative to the overall SME population. Internationalisation support that focuses only on leading firms cannot address the financing, compliance and channel constraints faced by a much broader population of enterprises.

Market structure and channel forms are changing at the same time. In 2025, China's trade with Belt and Road countries reached RMB 23.6 trillion, accounting for 51.9% of total foreign trade. Trade with ASEAN, the European Union, Latin America and Africa reached RMB 7.55 trillion, RMB 5.93 trillion, RMB 3.93 trillion and RMB 2.49 trillion, respectively (General Administration of Customs, 2026). Market diversification can reduce dependence on a single market, but it also increases institutional complexity. ASEAN markets involve regional supply-chain coordination and rules of origin; the European Union involves green, data and product-safety rules; and Latin American and African markets place greater weight on payment credit, exchange-rate risk, political risk and after-sales services. Cross-border e-commerce has also created new market entry points for SMEs. China's cross-border e-commerce imports and exports reached RMB 2.63 trillion in 2024, up 10.8% year on year, and RMB 2.75 trillion in 2025, 69.7% higher than in 2020 (China Government Network, 2025; General Administration of Customs, 2026). Yet the expansion of platform transactions does not equal the enhancement of firm-level internationalisation capability. Platforms control traffic, reviews, advertising, payments and account governance; firms without brand recognition, independent customer pools and after-sales systems may remain dependent on platforms. Born-global research also stresses that rapid internationalisation depends on innovation and organisational capability rather than channel convenience alone (Knight & Cavusgil, 2004).

Micro-level survey evidence further reveals the structural gap behind macro-level foreign-trade resilience. The World Bank (2025) shows that only 4% of firms in the China sample export directly, while the average time to clear exports through customs is three days, much better than the regional and income-group comparators. This suggests that customs efficiency is not the main constraint; the key issue is whether firms directly control overseas customers, market knowledge and brand premiums. If a firm participates in foreign trade only as an indirect supplier, its products may enter overseas markets, but customer relationships and market capabilities do not necessarily remain within the firm. The global environment also combines opportunities and constraints. UNCTAD estimates that global trade reached about USD 35 trillion in 2025, roughly USD 2.5 trillion more than in 2024 (UNCTAD, 2026). The WTO forecasts that world merchandise trade volume growth will fall from 4.6% in 2025 to 1.9% in 2026, and could fall to 1.4% under an energy-shock scenario. The share of trade covered by the most-favoured-nation principle also declined from about 80% in 2024 to about 72% in early 2026 (WTO, 2026). Global trade has not stopped growing, but rule fragmentation and uncertainty are raising entry costs for resource-constrained firms.

Taken together, these data show that facilitation and complication are occurring simultaneously. On the one hand, customs efficiency has improved, digital platforms have expanded and international market information is easier to obtain. On the other hand, carbon border adjustment, supply-chain due diligence, platform compliance, intellectual property protection and consumer-rights rules are becoming more demanding. Facilitation determines whether transactions can occur; complication tests whether firms can operate sustainably. For SMEs, the latter is often more difficult, because sustainable operation requires

organisational processes, compliance records, financing arrangements and after-sales systems. These are precisely the areas in which many firms are weakest. Foreign-trade aggregates therefore cannot be used as a substitute for firm performance. Growth in total trade may include many low-margin orders; growth in cross-border e-commerce may come with higher advertising costs and return rates; and market diversification may simultaneously introduce country risk and payment risk. Without indicators such as profit margins, payment periods, repurchase rates, return rates and financing coverage, it is difficult to judge whether firms have achieved high-quality internationalisation. Tables 1 to 8 jointly indicate that macro-level resilience must be converted into firm-level operating capability.

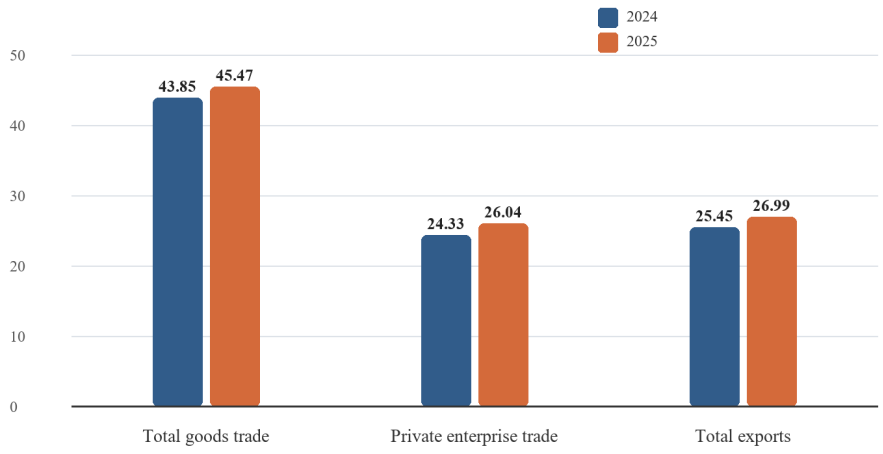
Chinese SMEs do not lack market entry points so much as they face a mismatch of capabilities. More than 60 million SMEs constitute a broad base of market actors, while only about 17,600 specialised and sophisticated 'little giant' firms have developed stronger international competitiveness through technology, specialisation and niche advantages (MIIT-related releases, 2026). This does not mean that ordinary SMEs have no possibility of internationalisation; rather, policy should not generalise the experience of a few leading firms to the entire population. Support should be stratified by technological capability, market capability and organisational capability. Market diversification also does not automatically reduce risk. Moving from European and American markets to ASEAN, Latin America or Africa may reduce dependence on one market, but it also introduces new tax, legal, payment, logistics and after-sales arrangements. Zaheer's (1995) liability of foreignness is relevant here: firms entering unfamiliar institutional environments must bear information, coordination and legitimacy costs. Cross-border e-commerce requires the same caution. Platforms lower entry thresholds, but they do not necessarily create customer assets. Oviatt and McDougall (1994) emphasise that international new ventures require resource-integration capability, which means that platform entry is not itself a competitive advantage. The decisive issue is whether a firm can transform entry into customer, brand and organisational capabilities.

Table 2. China's Foreign Trade and Private Enterprise Contribution, 2024-2025

Indicator	2024	2025	Change
Total goods trade	RMB 43.85 trillion; +5.0%	RMB 45.47 trillion; +3.8%	A new record high, with a slower growth rate.
Total exports	RMB 25.45 trillion; +7.1%	RMB 26.99 trillion; +6.1%	Exports remained the main source of resilience.
Total imports	About RMB 18.40 trillion	RMB 18.48 trillion; +0.5%	Import growth remained relatively weak.
Private enterprise trade	RMB 24.33 trillion; +8.8%	RMB 26.04 trillion; +7.1%	Private firms carried a larger share of growth responsibilities.
Private enterprises' share of total trade	55.5%	57.3%	Their position as a major trading body continued to rise.

Source: Compiled from public data released by the General Administration of Customs, China Government Network and China Economic Net.

Figure 1. China's Goods Trade Scale and Private Enterprise Contribution, 2024-2025



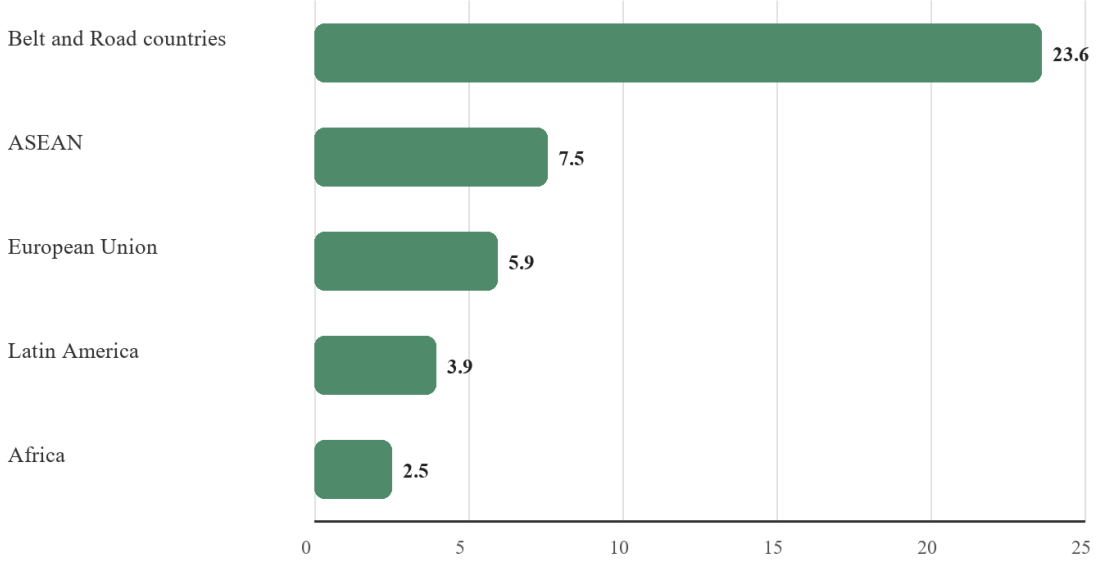
Source: Compiled from public data released by the General Administration of Customs and the State Council Information Office; unit: RMB trillion.

Table 3. Import and Export Scale of Key Markets in 2025

Market	Scale	Year-on-year growth	Risk implication
Belt and Road countries	RMB 23.6 trillion	+6.3%	Market expansion, but larger cross-country institutional differences.
ASEAN	RMB 7.55 trillion	+8.0%	Supply-chain cooperation and localisation requirements coexist.
European Union	RMB 5.93 trillion	+6.0%	Dense green, data and product-safety rules.
Latin America	RMB 3.93 trillion	+6.5%	Higher pressure from exchange rates, credit and logistics cycles.
Africa	RMB 2.49 trillion	+18.4%	High growth, but prominent after-sales and payment risks.

Source: Compiled from the General Administration of Customs' full-year 2025 foreign trade release.

Figure 2. Import and Export Scale of Key Markets in 2025



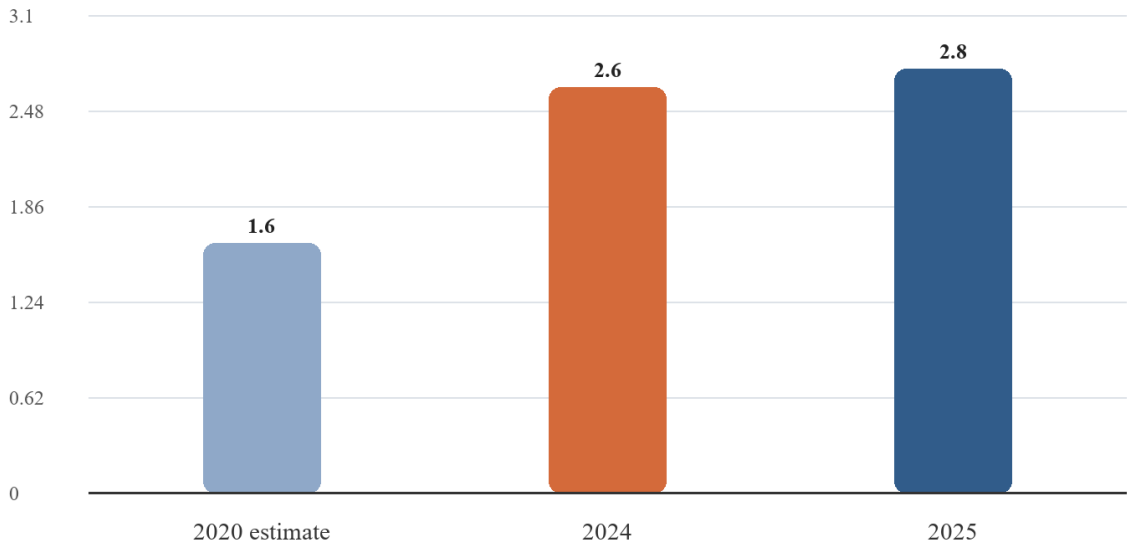
Source: Compiled from the General Administration of Customs' full-year 2025 foreign trade release; unit: RMB trillion.

Table 4. Changes in Cross-Border E-Commerce Import and Export Scale

Year	Scale	Change	Implication
2020	About RMB 1.62 trillion	Derived from the statement that 2025 was 69.7% higher than 2020	The digital-channel base was already substantial.
2024	RMB 2.63 trillion	+10.8%	Platform-based entry continued to expand.
2025	RMB 2.75 trillion	+69.7% compared with 2020	Transaction growth does not equal the accumulation of customer assets.

Source: Compiled from public data released by the General Administration of Customs and China Government Network; the 2020 value is derived from 2.75/1.697.

Figure 3. Changes in China's Cross-Border E-Commerce Import and Export Scale



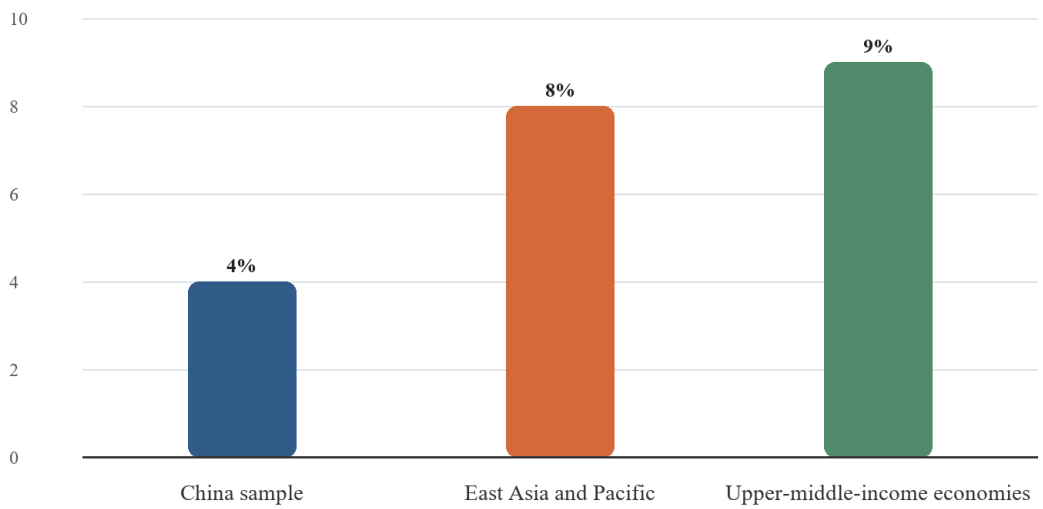
Source: Compiled from public data released by the General Administration of Customs and China Government Network; the 2020 value is derived from 2.75/1.697; unit: RMB trillion.

Table 5. Micro-Level Gaps in the World Bank China Enterprise Survey

Indicator	China 2024	East Asia and Pacific	Upper-middle-income economies	Implication
Share of firms exporting directly	4%	8%	9%	Direct firm-level internationalisation remains low.
Average days to clear exports through customs	3 days	9 days	10 days	Customs clearance is not the main bottleneck.
Average days to clear imports through customs	11 days	7 days	11 days	Import lead time still affects production scheduling.
Small firms with bank loans	22%	-	-	Small firms face evident financing disadvantages.
Medium-sized/large firms with bank loans	46%/46%	-	-	Access to credit is strongly stratified by size.

Source: World Bank Enterprise Surveys, China 2024 Country Profile.

Figure 4. International Comparison of Firms' Direct Export Participation



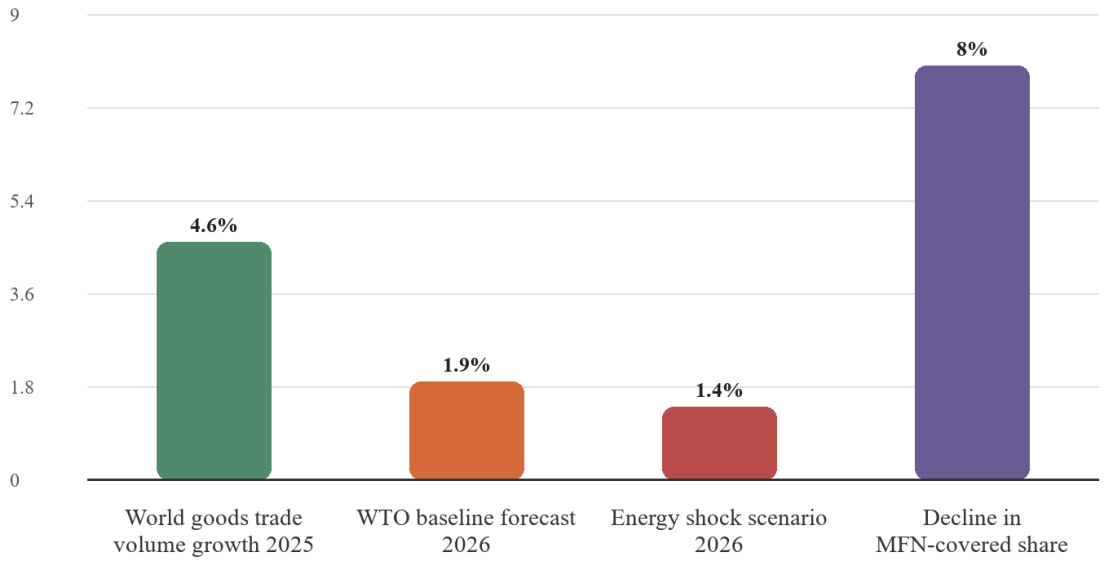
Source: World Bank Enterprise Surveys, China 2024 Country Profile.

Table 6. Global Trade and Rules Environment

Variable	Value or node	Interpretation
Global trade scale	About USD 35 trillion in 2025	The aggregate volume was still expanding.
Increase in global trade	About USD 2.5 trillion more than in 2024	Market opportunities still existed.
World merchandise trade volume growth	4.6% in 2025; WTO baseline forecast of 1.9% in 2026	Growth was expected to slow.
Energy shock scenario	May fall to 1.4% in 2026	External shocks could still compress orders.
MFN-covered trade share	About 80% in 2024; about 72% in early 2026	Rule fragmentation was becoming more visible.

Source: UNCTAD Global Trade Update 2026 and WTO Global Trade Outlook and Statistics 2026.

Figure 5. Global Trade Growth and Rule Fragmentation Pressure



Source: WTO Global Trade Outlook and Statistics, March 2026; the decline in MFN-covered trade share is calculated from 80% to 72%.

Table 7. Timeline of Green and Supply-Chain Rules

Rule	Application node	Scope	Pressure on SMEs
EU CBAM	Definitive period from 1 January 2026	Carbon-emission reporting and certificate arrangements	Pressure from carbon data, authorised reporting and cost pass-through.
EUDR	Large and medium-sized firms: 30 December 2025; small and micro firms: 30 June 2026	Due diligence for deforestation-related products	Supply-chain traceability and material documentation requirements.
EU supply-chain and product-safety rules	Progressively strengthened	Product safety, ESG and consumer protection	Raises documentation, legal and after-sales costs.

Source: Compiled from public information released by the European Commission and the Council of the European Union.

3. Formation Mechanisms

Low-price entry is a common initial route for many SMEs entering international markets, but it can easily lead to value-chain lock-in. When firms lack brands, channels and customer trust, they often rely on low prices to secure overseas orders. Once many firms in the same industrial cluster cut prices simultaneously, low prices shift from an individual competitive tactic to a buyer's reference point for bargaining, weakening the premium-earning capacity of the entire chain. Gereffi et al. (2005)

argue that buyer-driven global value chains restrict suppliers' upgrading space. In SME export settings, this buyer power appears in asymmetric price benchmarks, inspection standards, payment terms and responsibility allocation. If firms fail to include certification, after-sales service, exchange-rate exposure, returns, insurance and intellectual property costs in their quotations, growth in export scale may weaken profit stability. Even firms with higher quality, more complete compliance documents and more stable delivery may not receive an adequate premium.

Compliance, financing and payment-period constraints are becoming core pressures in SME internationalisation. The EU Carbon Border Adjustment Mechanism (CBAM) entered its definitive period on 1 January 2026, and the EU Deforestation Regulation (EUDR) has also specified the application timetable for large, medium-sized, small and micro firms (European Commission, 2026; Council of the European Union, 2025). Such rules require firms to provide continuing data on carbon emissions, raw-material origins, supply-chain due diligence and production processes. Compliance costs are shifting from one-off certification to continuous data governance. Large firms can spread these costs through specialised departments and large orders. SMEs, by contrast, have more difficulty absorbing fixed compliance costs and are more vulnerable to incomplete materials, misunderstandings of rules or unstable service-provider quality. Financing constraints amplify this pressure. The World Bank (2025) reports that 22% of small firms in the China sample have bank loans, compared with 46% of both medium-sized and large firms. ADB (2025) estimates the global trade-finance gap at about USD 2.5 trillion. Cross-border orders often require firms to pay in advance for procurement, production, logistics, certification and platform fees, while payment collection may lag. If financial institutions cannot identify real trade through customs declarations, tax invoices, logistics traces, credit insurance, platform receipts and historical repayment records, firms cannot easily convert orders into financing capacity.

Platform-based internationalisation lowers the threshold of market entry, but it also concentrates the power to set operational rules. Cross-border e-commerce platforms help firms reach overseas consumers, yet they also control traffic allocation, advertising auctions, rating systems, return policies, payment settlement and account sanctions. Platforms are not inherently negative; they reduce search and transaction costs. However, platform rules are not the same as firm strategy, and platform sales should not be mistaken for brand assets. Firms without diversified channels or independent customer accumulation may see their sales stability directly affected by platform rule changes. Oviatt and McDougall (1994) emphasise that international new ventures need to integrate resources and networks, while Freeman et al. (2006) show that small firms can use networks to overcome internationalisation constraints. Platforms can therefore serve as network entry points, but they cannot replace a firm's own customer assets, after-sales systems or organisational capabilities. Brand capability should also be understood rigorously. It is not merely packaging, advertising or an English name; it is the credible expectation that a firm will provide stable quality, reliable delivery, traceable after-sales service, complete compliance materials and standardised dispute handling. Cavusgil and Knight's (2015) work on capabilities and resource allocation likewise shows that a brand is, in essence, the external expression of organisational capability.

Insufficient organisational capability prevents internationalisation experience from accumulating and magnifies sector-specific pressures. Many SMEs still rely on the personal judgement of owners and the experience of sales staff, while lacking basic management tools such as customer credit sheets, order profit sheets, compliance-document files and after-sales issue records. Johanson and Vahlne (1977) stress the accumulation of experiential knowledge, but experience that is not transformed into processes cannot become firm capability. Leonidou (2004) also identifies insufficient managerial resources as a barrier to small-firm export development. Export departments know customer commitments, production departments know quality risks, finance departments know payment pressures, and after-sales staff know product problems. If this information is not recorded and reviewed centrally, it cannot inform the next round of quotations, contract review or market selection. Sectoral differences also matter. Textile and apparel firms mainly face returns, inventory and labour-compliance pressures; furniture and home-product firms face bulky logistics, overseas-warehouse turnover and after-sales costs; automotive and motorcycle parts firms must handle technical adaptation and product liability; consumer electronics firms are affected by platform complaints, advertising costs and data privacy; and new-energy supporting firms face carbon footprints, trade remedies and local-content rules. The internationalisation difficulties of SMEs are therefore not the result of a single external shock, but of

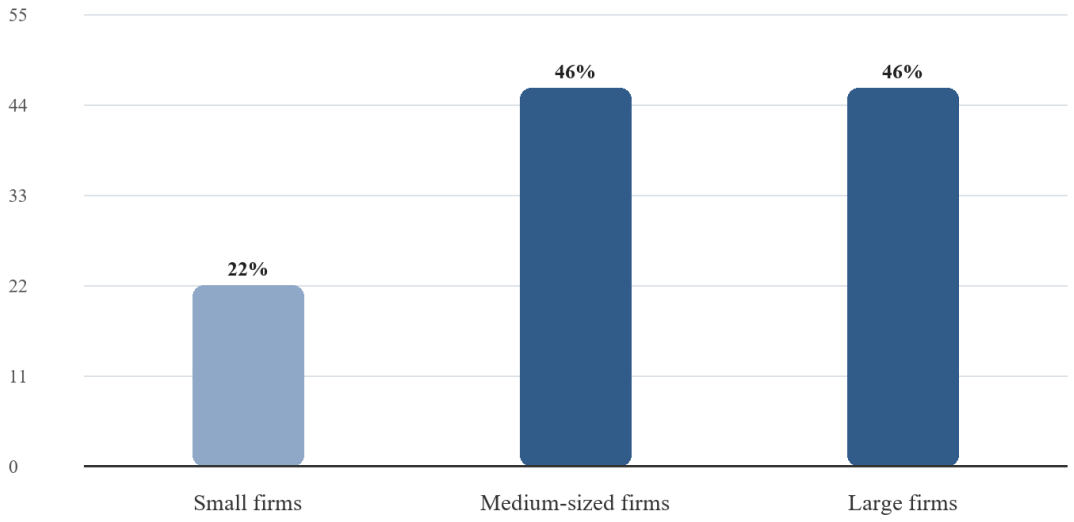
low-price competition, compliance costs, financing shortages, platform dependence and weak organisational capability acting together. Policy that provides only uniform training, without sectoral indicators, templates and data tools, will find it difficult to reduce firms' real operating costs.

Table 8. Financing and Cash-Flow Pressure Data

Source	Indicator	Value	Meaning for SMEs
World Bank	Small firms with bank loans	22%	Weak order-financing capacity.
World Bank	Medium-sized/large firms with bank loans	46%/46%	Clear size-based stratification.
ADB	Global trade finance gap	About USD 2.5 trillion	Trade-finance shortage remains a global problem.
Public releases by the People's Bank of China	Balance of inclusive loans to small and micro firms	About RMB 32.93 trillion	The aggregate supply has improved, but foreign-trade scenarios still require order-based recognition.

Source: Compiled from public data released by the World Bank, ADB, China Government Network and the People's Bank of China.

Figure 6. Share of Chinese Sample Firms with Bank Loans by Size



Source: World Bank Enterprise Surveys, China 2024 Country Profile.

Table 9. Typical Sectoral Constraints and Data-Based Observation Indicators

Sector	Main internationalisation mode	Key cost	Suggested tracking indicator
Textiles and apparel	OEM production and platform retailing	Returns, labour compliance and inventory	Return rate, repurchase rate and unit gross margin.
Furniture and home products	Trade-fair orders and overseas warehouses	Bulky logistics, after-sales service and slow-moving inventory	Inventory turnover days and after-sales response time.
Automotive and motorcycle parts	B2B and aftermarket sales	Adaptation data and liability risk	Complaint rate and certification coverage.
Consumer electronics	Platform stores and independent brands	Advertising costs, product safety and privacy rules	Advertising-to-sales ratio and account health.
New-energy supporting products	Components, equipment and engineering support	Carbon footprint and trade-remedy exposure	Carbon-data completeness and compliance audit frequency.

Source: Summarised by the author from public rules, sectoral internationalisation patterns and the international business literature.

4. Governance Pathways

Policy for SME internationalisation should shift from scale orientation to capability orientation. Export value, the number of trade fairs attended, the number of cross-border e-commerce stores, overseas-warehouse area and training sessions can indicate the intensity of market entry, but they do not show whether firms earn profits, collect payments on time, possess compliance capability, retain customers or sustain after-sales service. The OECD (2009) notes that barriers to SME internationalisation are not confined to market entry; they include information, finance, managerial capability and external procedures. Public services should therefore be closer to firms' operating scenes. For firms entering the EU market, public services should provide product-code-based rule lists, certification pathways, document templates, contract clauses, service-provider evaluations and cost ranges. For platform sellers, they should offer tools for intellectual property searches, account health, product safety and tax compliance. For firms in industrial clusters, shared testing, overseas after-sales services, dispute cases and customer-credit databases should be developed. These tools can reduce fixed costs more effectively than general lectures and are better suited to the reality that SMEs have limited resources and insufficient capacity to interpret rules.

Trade finance and industrial-cluster governance should also move from general support to scenario-based support. Financial institutions should not rely only on collateral logic. Subject to compliance requirements, they should integrate orders, customs declarations, tax invoices, logistics information, credit insurance, platform receipts and collection records to develop small, revolving trade-finance products that match payment cycles. The trade-finance gap identified by ADB (2025) shows that financing shortages are not only a matter of funding supply, but also of identifying real trade and sharing risk. Industrial clusters should shift from low-price agglomerations to capability communities. They can convert fixed costs that individual firms cannot bear - such as compliance, certification, after-sales service and dispute handling - into public-service costs through shared compliance databases, shared certification resources, shared overseas after-sales nodes and joint dispute-resolution mechanisms. Global value-chain research emphasises the effect of network relations and governance structures on firms' upgrading space (Gereffi et al., 2005); public services in industrial clusters are therefore an important path for improving SMEs' value-chain positions. Local governments, industry associations, financial institutions and platform firms should also coordinate. Industry associations can accumulate certification procedures, service-provider evaluations and contract templates; platform firms can use transaction, return, complaint and account-risk data to provide compliance warnings; financial institutions can improve their ability to identify real trade; and government departments can coordinate rules and provide public resources.

Firms themselves also need to establish a minimum governance structure for internationalisation. Owners should define market-entry choices and risk limits; export departments should manage customers and channels; finance departments should manage exchange rates, payment terms and financing; production departments should manage quality and batch traceability; after-sales departments should close complaint loops; and external legal and certification agencies should provide specialised support. The structure need not be large, but responsibility boundaries must be clear. Firms should also establish customer segmentation, quotation cost models, contract review, compliance files and order-review mechanisms so that they are less dependent on individual experience and ad hoc responses. Evaluation systems should enter the firm's accounts directly. Rising export value with falling net profit margins should not be regarded as high-quality internationalisation; rising GMV with an uncontrolled advertising-to-sales ratio should not be regarded as brand improvement; and expansion of overseas-warehouse area with worsening inventory turnover should not be regarded as stronger service capacity. More appropriate indicators include export net profit margin, collection cycle, certification pass rate, credit-insurance coverage, advertising-to-sales ratio, repurchase rate, return-processing time and after-sales complaint rate. Only when these indicators improve continuously can internationalisation become an enhancement of overall firm governance rather than merely a sales task for the business department.

Table 10. Governance Shift from Scale Indicators to Capability Indicators

Old indicator	Problem	New indicator	Use
Export value	Conceals profit and payment cycles	Export net profit margin and collection cycle	Measuring operating quality.
Number of stores and GMV	Conceals platform dependence	Advertising-to-sales ratio, repurchase rate and account health	Measuring channel quality.
Number of trade fairs	Difficult to evaluate subsequent transactions	Customer repurchase and sample-to-order conversion rate	Measuring market-development quality.
Overseas warehouse area	Conceals inventory risk	Turnover days and return-processing time	Measuring service quality.
Training sessions	Can become formalistic	Template use and certification pass rate	Measuring public-service effectiveness.

Source: Summarised by the author from OECD research on SME internationalisation and the data analysis in this article.

5. Conclusion

The central difficulty in the internationalisation of Chinese SMEs is not a shortage of overseas market opportunities, but a structural mismatch between firm capabilities and external rules. Total foreign-trade volume, the share of private enterprises and the scale of cross-border e-commerce show that SMEs remain an important force in China's foreign-trade expansion. Yet data from the World Bank, WTO, UNCTAD and ADB also show that insufficient direct export participation, trade-finance constraints, slower global trade growth and rule fragmentation are continuously raising the costs of entering and operating in overseas markets (World Bank, 2025; WTO, 2026; UNCTAD, 2026; ADB, 2025). SME internationalisation should therefore not stop at 'products going global'; it should move towards 'capabilities going global'. Low-price orders, platform traffic and entry subsidies may bring short-term orders, but they rarely produce stable returns. Long-term competitiveness rests on the institutionalisation of quality control, compliance response, credit management, order financing, customer accumulation, after-sales service and organisational processes. As global trade rules become increasingly detailed, firms no longer compete only on price; they compete on delivery stability, data completeness, compliance credibility and the ability to control transaction risk.

A more robust form of SME internationalisation should be assessed through sustainable operation. Whether a firm can secure stable repurchases, include compliance costs in its quotations, improve cash flow through credit insurance and order financing, accumulate customer assets beyond platforms, and use after-sales data to improve products says more about the quality of internationalisation than the size of a single order. Public policy should accordingly reduce its dependence on visible indicators such as trade fairs, stores and transaction value, and turn instead to sectoral compliance tools, scenario-based trade finance, public services for industrial clusters and data-based risk governance. Firms should incorporate customer, order, compliance, financing and after-sales information into their internal governance structures. Chinese SMEs have advantages in matching efficiency, manufacturing experience and market responsiveness, but these strengths can withstand rule changes and buyer-governance pressure only when they are combined with compliance, finance, brand and organisational capabilities. Only when profit, payment collection, repurchase, compliance and after-sales indicators improve continuously can high-quality foreign-trade development move beyond macro-level scale growth and become long-term competitiveness at the micro level of SMEs.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Asian Development Bank. (2025). ADB Global Trade Finance Gap Survey. https://www.adb.org/publications/adb-global-trade-finance-gap-survey

[2] Cavusgil, S. T., & Knight, G. (2015). The born global firm: An entrepreneurial and capabilities perspective on early and rapid internationalization. Journal of International Business Studies, 46(1), 3-16.

[3] Council of the European Union. (2025). Deforestation: Council signs off targeted revision to simplify and postpone the regulation. https://www.consilium.europa.eu/en/press/press-releases/2025/12/18/deforestation-council-signs-off-targeted-revision-to-simplify-and-postpone-the-regulation/

[4] Dunning, J. H. (1988). The eclectic paradigm of international production: A restatement and some possible extensions. Journal of International Business Studies, 19(1), 1-31.

[5] European Commission. (2026). Carbon Border Adjustment Mechanism. https://taxation-customs.ec.europa.eu/carbon-border-adjustment-mechanism_en

[6] Freeman, S., Edwards, R., & Schroder, B. (2006). How smaller born-global firms use networks and alliances to overcome constraints to rapid internationalization. Journal of International Marketing, 14(3), 33-63.

[7] Gereffi, G., Humphrey, J., & Sturgeon, T. (2005). The governance of global value chains. Review of International Political Economy, 12(1), 78-104.

[8] Johanson, J., & Vahlne, J. E. (1977). The internationalization process of the firm. Journal of International Business Studies, 8(1), 23-32.

[9] Knight, G. A., & Cavusgil, S. T. (2004). Innovation, organizational capabilities, and the born-global firm. Journal of International Business Studies, 35(2), 124-141.

[10] Leonidou, L. C. (2004). An analysis of the barriers hindering small business export development. Journal of Small Business Management, 42(3), 279-302.

[11] Lu, J. W., & Beamish, P. W. (2001). The internationalization and performance of SMEs. Strategic Management Journal, 22(6-7), 565-586.

[12] OECD. (2009). Top barriers and drivers to SME internationalisation. OECD Working Party on SMEs and Entrepreneurship.

[13] Oviatt, B. M., & McDougall, P. P. (1994). Toward a theory of international new ventures. Journal of International Business Studies, 25(1), 45-64.

[14] UN Trade and Development. (2026). Global Trade Update: April 2026. https://unctad.org/publication/global-trade-update-april-2026-global-trade-growth-continues-fragility-rises

[15] World Bank. (2025). Enterprise Surveys: China 2024 Country Profile. https://www.enterprisesurveys.org/content/dam/enterprisesurveys/documents/country/China-2024.pdf

[16] World Trade Organization. (2026). Global Trade Outlook and Statistics: March 2026. https://www.wto.org/english/res_e/publications_e/gtos0326_e.htm

[17] Zaheer, S. (1995). Overcoming the liability of foreignness. Academy of Management Journal, 38(2), 341-363.

[18] China Government Network. (2025). China's foreign trade exceeded RMB 43 trillion in 2024, reaching a new record high. https://www.gov.cn/lianbo/bumen/202501/content_6998197.htm

[19] China Government Network. (2025). Chart: China's cross-border e-commerce imports and exports grew by 10.8% in 2024. https://www.gov.cn/zhengce/jiedu/tujie/202501/content_6998286.htm

[20] MIIT-related releases. (2026). China has cultivated more than 140,000 specialised and sophisticated SMEs. https://www.gov.cn/yaowen/shipin/202501/content_6996829.htm

[21] General Administration of Customs. (2026). China's total import and export value exceeded RMB 45 trillion in 2025, reaching a historical high. https://www.ce.cn/xwzx/gnsz/gdxw/202601/t20260114_2701371.shtml

[22] General Administration of Customs. (2026). Cross-border e-commerce imports and exports reached RMB 2.75 trillion in 2025, 69.7% higher than in 2020. https://www.ce.cn/xwzx/gnsz/gdxw/202601/t20260114_2700917.shtml

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

The Impact of Women's Leadership on Firm Performance

Shujun Chen*

Sun Life Everbright Life Insurance Co., Ltd., Shanghai Branch, Shanghai, 200082, China

Corresponding author: Shujun Chen, 15021122158@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This systematic review integrates interdisciplinary literature on the impact of women's leadership on firm performance, based on 40 core articles (AJG 2024 4*/4) and 10 supplementary landmark studies. Four theoretical perspectives are identified: social roles, resource supplementation, contextual factors, and behavioral practices. The review reveals that the influence of women's leadership is not a simple positive or negative effect, but a complex process dynamically regulated by bias challenges, resource complementarity, situational conditions, and strategic behaviors. Core debates center on the shift from simple judgments to situational dependence, from stereotype constraints to strategic integration of agency and communion, and from numerical representation to qualitative capital. Key gaps are identified, including the lack of multi-level interaction theories, sample biases toward developed countries, and endogeneity issues in causal identification. Future research should focus on multi-level interactive processes and expand exploration in Global South contexts such as China.

Keywords: Women's Leadership; Firm Performance; Gender Diversity; Role Congruity Theory; Resource-Based View; Situational Factors; Behavioral Strategies; Systematic Review

Published: Jun 10, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1494

1. Introduction

1.1 Description of Women's Leadership and Firm Performance

Women's leadership is typically characterized by the ability to balance stakeholders, care, empathetic listening, and continuous self-reflection. These traits contribute to strengthening team cohesion during times of crisis, thereby building up key momentum for ensuring and enhancing firm performance, which includes a multi-dimensional integrated concept of financial results, strategic execution, and organizational effectiveness. At the same time, a diverse perspective at the organizational level helps inspire innovative solutions, shape a more growth-oriented organizational atmosphere, and drive a smooth and sustainable process of reformation within the firm.

1.2 Review Significance

Research on the impact of women's leadership on firm performance spans multiple disciplines including strategy, organizational behavior, human resource management, economics, entrepreneurship, psychology, sociology, and business history. The perspectives and conclusions of existing research are mostly fragmented. Therefore, this review aims to systematically analyze, critically evaluate and integrate the existing interdisciplinary literature to clarify the core theoretical context of the field, reveal research contradictions and gaps, and thus point out the direction for future research. An in-depth exploration of this topic is not only about enhancing firm competitiveness and organizational resilience, but also involving deeper themes such as social structure optimization and sustainable development.

1.3 Review Purpose and Scope

This review aims to integrate fragmented existing knowledge by comparing and synthesizing different theoretical lineages to reveal their intrinsic connections and evolutionary logic, thereby constructing a more systematic and dynamic understanding framework. The specific objectives include: (1) a systematic review of the formation and evolution of theories related to the impact of women's leadership on firm performance; (2) Based on existing literature, deeply summarize and analyze the interaction of psychological mechanisms, behavioral strategies, and situational factors behind it; (3) Evaluate the development of existing theories from both theoretical and empirical perspectives, accurately identify theoretical and methodological gaps, and clarify future research directions.

1.4 Review Structure and Main Findings

This review reveals that the impact of women's leadership on firm performance is a complex and dynamic process interwoven with multiple mechanisms and scenarios. Research in this field has evolved from early demonstrations of gender role differences to the complementary value of resources brought about by gender diversity, and further to the analysis of the action mechanism of specific situational factors and behavioral practices. However, the full exertion of women's leadership still faces challenges such as the "glass ceiling" and underrepresentation in high-growth industries. Existing research findings are also contradictory, requiring an in-depth analysis of the complex interactions of social development, psychological mechanisms, behavioral strategies, and situational conditions. It is necessary to go beyond a simple description of correlation. At the same time, the existing research still has significant limitations, such as sample limitations leading to insufficient connections to the broader field, and failure to fully capture the dynamic interaction processes at multiple levels (individual, team, organization). Finally, this review aims to effectively outline the linkages and tensions of different theories, systematically reveal research gaps, and lay the foundation for constructing an integrated research framework that is more explanatory and sensitive.

2. Method

2.1 Type of Review

This systematic review aims to conduct a comprehensive search, screening, evaluation and integration of high-quality core literature on the impact of women's leadership on firm performance, by using a pre-defined, clear and repeatable methods.

2.2 Search Strategy

To ensure as comprehensive coverage as possible of all relevant high-quality academic articles and to reach as objective and reliable conclusions as possible, a systematic and comprehensive search strategy was employed in this study. According to the authoritative "Journal Quality List" and the relevant subject area, journal articles with AJG rated of 4 and 4* were prioritized for search to quickly identify high-impact core literature.

2.2.1 Data Inclusion and Exclusion

Subject Areas: This review focuses on journals in the areas of business history, economics, entrepreneurship, general and strategy, organizational behavior, human resource management, psychology, and sociology. According to "Journal Quality List", journals with AJG rated of 4 and 4* in these areas were selected. A total of 95 ISSN numbers were obtained as the basis for journal search.

Keywords: A total of 532 related articles were initially obtained by searching 95 journal databases in sequence through Stirling Library Search/Journals/Journal Search, using the keywords "women* leader*".

Publication year, language, title, abstract: To ensure the timeliness of the articles, the publication year is refined from 2015 to the present (the last ten years). The language is refined to English. Preliminary screening through quick reading of titles and abstracts will exclude articles that are overly focused on regional, racial, religious, political, military, specific industry, and do not substantially address the core theories and concepts of how women's leadership affects firm performance. After this step, the number of articles is refined from 532 to 40 high-quality articles downloaded to the local folder for full-text reading and review.

Table 1. Data statistics of the literature search and screening process.

Subject areas	Number of journals (AJG 2024 4*-4)	Initial number of articles (keyword search)	Number of articles after filtering
Bus Hist	3	11	1
Economics	29	21	0
Entrep	2	6	1
Gen & Strat	24	82	9
OB/OS/HRM	15	204	25
Psychology	14	93	4
Sociology	8	115	0
Total	95	532	40

2.2.2 Data Extraction

Key assessment: Extract key points from the last 40 articles and input them into an Excel sheet for group management and quick review. The specific contents are as follows:

(1) Basic information: Title, abstract, author, publication year.

(2) Chapter structure: The main subtopics.

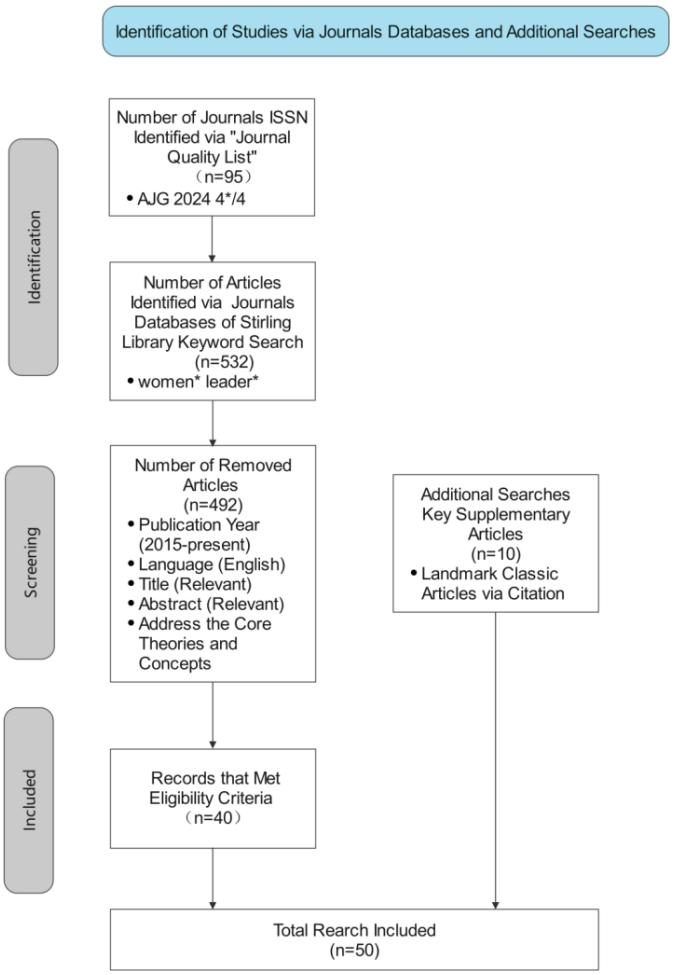
(3) Research methods: Empirical analysis, case studies, meta-analysis, etc.

(4) Theoretical perspective: The core theoretical framework.

(5) Key conclusions: The main findings.

(6) Contradictions or gaps: Disputes or recommendations for future research.

Figure 3. Search process and results



Supplementary materials: During the intensive reading of the 40 articles, 10 landmark classic articles were traced and supplemented through its citation. It is not only broadening the theoretical depth and historical context of this literature review, but also ensuring the comprehensiveness of the analysis. Such as "role consistency theory" (Eagly & Karau, 2002), "glass ceiling theory" (Dreher, 2003), "glass cliff phenomenon" (Ryan & Haslam, 2005), "gender stereotypes" (Heilman, 2012; Heilman & Okimoto, 2007; Rudman & Phelan, 2008; Sanchez & Lehnert, 2019), " agency and communion" (Zheng, Surgevil, et al., 2018), "crisis advantage" (Post et al., 2019), "TMT diversity" (Dezso & Ross, 2012).

2.3 Search Process and Result Overview

The process begins with the source of high-quality literature, that is, by screening out journals with AJG rated of 4 and 4* in specific interdisciplinary areas based on the authoritative "Journal Quality List". It is to ensure the authority and high relevance of the start for the literature review.

Ultimately, a high-quality literature pool consisting of 40 core articles and 10 key supplementary articles was formed, through the initial search and acquisition, strict inclusion and exclusion, structured information extraction and management. It lays a solid foundation for the subsequent objective and reliable conclusions.

3. Literature Review

3.1 Literature Classification: Theories from Multiple Perspectives

Through the sorting and analysis of the extracted information, we can find that the existing research on the impact of women's leadership on firm performance can be roughly classified into four categories based on their core theoretical perspectives: the perspective based on social roles, resource supplementation, situational factors, and behavioral practice.

3.1.1 Perspective Based on Social Roles

The core argument of this perspective is centered around the theory of role consistency (Eagly & Karau, 2002) and the theory of legitimacy (Vial et al., 2016). It is argued that gender role stereotypes (Heilman & Okimoto, 2007) may lead to biased evaluations, thereby hindering women leaders from driving performance improvement.

Women leaders face a "double bind". If they display more agentic (such as being capable, ambitious, and competitive), they may be perceived as lacking in community and receive negative evaluations; if they show community (such as being caring, cooperative, relationship-oriented and fair), they may be regarded as lacking in leadership and be seen as incompetent (Rudman & Phelan, 2008). Employees generally prefer to work for man bosses, and this preference directly affects the working environment and support for women leaders, which in turn influences decision-making and behavior (Powell & Anthony Butterfield, 2015). Women-founded enterprises receive less labor input from employees at the same salary level, which directly affects their strategic execution ability (Kacperczyk et al., 2023).

Leadership role perceptions, gender biases and restraints have changed, however. (Schaumberg & Flynn, 2017) indicate that the trait of "self-reliance" can enhance the evaluation of women leaders because they are often regarded as having the same level of ability as men while also being considered more approachable. Gender bias is relatively weaker in unconventional (such as innovative, risky) projects (Parker et al., 2020). Seeing a woman in a leadership role can communicate that an organization can be trusted (Joshi & Diekman, 2022).

In the current era when there is a strong advocacy for enhancing the participation of women in decision-making positions, it is not only necessary to pay attention to external biases, but also to focus on enhancing the inner well-being of women leaders. While leadership and related resources such as power and status are often depicted as positive experiences, taking on a leadership role often brings more negative subjective emotional experiences and emotional exhaustion to women (Ong, 2022).

3.1.2 Perspective Based on Resource Supplement

This perspective views women leaders as an important strategic resource for organizations. They can bring diversity in terms of information and social aspects to the top management team (TMT), thereby enriching the decision-making perspectives and behavioral patterns of the management, and exerting a positive role model and motivational effect on middle-level women managers. (Dezso & Ross, 2012).

Multiple studies have supported the resource value of gender diversity in TMT. The meta-analysis (Post & Byron, 2015)

confirms that women directors' professional knowledge and rich experience can enable the board of directors to participate more deeply in the strategic planning process and provide professional advice. (Dai et al., 2019) provide empirical support that adding women management members in new ventures teams, especially high-tech ones, has a positive impact on innovation performance through expanded knowledge differentiation and integration, as well as diversification of information processing and managerial styles. (Schultheiss, 2021) calls f or attention to the design at the top, arguing that women leaders have created a positive demonstration effect within the firm and passed on its influence. It exerts a profound "trickle-down effect" on the overall organizational culture. (Kirsch, 2018) suggests that increasing the proportion of women on the board of directors not only helps promote gender equality in the economic field, but also aligns with organizational interests. The decision-making mechanism has been optimized and the benefits have been extended to society and ethical aspects of corporate behavior, which is a best practice to follow in corporate governance. (Sieweke et al., 2023) adopt an instrumental variable (IV) design to estimate the impact of gender diversity of TMT. Experiments have demonstrated its impact on the company's profitability, liquidity and growth.

Nevertheless, the effect of gender diversity is also influenced by power structure and contextual factors. Board decisions follow the position of the relatively more influential "boy's club", but reverse if the chair is a woman (Sidhu et al., 2021). A high level of women leaders in the senior management team and the board of directors will bring about more organizational innovation, which will ultimately enhance the performance of the enterprise (Wu et al., 2022).

The positive impact of women leaders on performance remains constrained by persistent biases and structural barriers. Women directors may have been appointed initially due to institutional pressure (Kirsch, 2018). However, the advantages of external quota systems and others on the board's response to gender bias remains limited (Knippen et al., 2019). Although women directors help maximize the practices and policies in product innovation and governance (Glass & Cook, 2018). They are also often promoted to high-risk leadership positions, but lack the support or authority needed to achieve their strategic goals (Glass & Cook, 2016).

From the perspective of political capital, the sources of influence for women directors are specifically classified into eight categories: economic, cultural, knowledge, social, symbolic, reputational, organizational, and institutional. When the proportion of women directors reaches a critical mass (roughly 30-33%), the accumulation of their symbolic capital will endow them greater say and influence. The influence of women directors is mainly transmitted through board functions (monitoring and advising). The monitoring function may manifest as stricter regulation of R&D investment, while the advising function is reflected in providing unique advice for innovation strategies (Yang & Konrad, 2025).

3.1.3 Perspective Based on Contextual Factors

This perspective emphasizes the core regulatory role of both internal and external situational factors. It is not universal but strongly influenced by specific conditions such as national institutions (gender quotas), organizational culture (gender equality, board support), project nature (non-traditional), task characteristics (high risk), team composition (coordination needs), as well as the characteristics and transition context of previous leaders.

Early research suggests that women may face "glass cliff" (Ryan & Haslam, 2005) after breaking through the career "glass ceiling" (Dreher, 2003). That is to say, they are often appointed to risky and unsteady management positions, and it is regulated by multiple situational factors (Ryan et al., 2016). Women leaders often have more trust advantages in interpersonal emotion management, communication and relational abilities during organizational crises (Post et al., 2019). Empirical study supports the view that female leaders are more favored during crisis because they are more empathetic and confident (Sergent & Stajkovic, 2020). Furthermore, gender bias may be weakened when the projects led by women are unconventional (innovative and risky), as they move away from male-centric prototypes (Parker et al., 2020).

Research shows that in teams with high coordination requirements (such as large size, functional diversity, or geographical dispersion), women leaders tend to enhance the team's cohesion, cultivate the awareness of collaborative learning, and demonstrate a more proactive willingness to communicate. Their democratic and relationship-oriented leadership style contributes to information sharing, psychological safety, and knowledge integration (Post, 2015). Badura et al. (2018) suggest that gender differences can be reduced in business environments, long-term interactions, and situations with high social

complexity.

In the context of CEO succession, women CEOs often face the threat of stereotypes, which may disrupt the long-term innovation strategies. However, its positive effect is more pronounced in organizational environments supported by women executives (Dwivedi et al., 2021). Women CEOs have encompassed key entrepreneurial roles in emerging markets, such as the Colombian women entrepreneurs (Davila & Lluch, 2022). In a highly scrutinized environment, women CEOs adjust their information processing and risk propensity, thereby influencing innovative strategies such as mergers and acquisitions (Gamache et al., 2023).

At the institutional level, gender quota laws tend to enhance the information "bridges" of women directors in the directors' network, while corporate governance codes help facilitate their access to information (Mateos De Cabo et al., 2022). Studies of diversity management practices in multinational enterprises also suggest that formal social institutions (gender-equality) have moderating effects on women leadership abilities (Terpstra-tong et al., 2025). Furthermore, informal cultural values, incentives, mentorship can encourage and enhance women's aspirations and job satisfaction in competing for leadership (Sanchez & Lehnert, 2019). (Wille et al. (2018) also suggest that higher performance among women-led firms reflects a fair organizational culture, and actively fulfilling corporate social responsibilities.

Women leaders can also drive structural change within the organization. The trickle-down effect can drive change in two ways. (1) agency perspective (advocates, mentors, and role models that bring more management opportunities to other women) (2) signaling theory (having diversity on boards may signal to investors a company's social values) (Stainback et al., 2024). The pipeline effect echoes the "bottom-up" theory, which suggests that in innovation-oriented or labor-intensive enterprises, women leaders can activate grassroots execution by fostering employee participation, and empowering employee voice (Jeong et al., 2025). It is a new process explanation for the impact of women leadership on performance.

3.1.4 Perspective Based on Behavioral Practices

This perspective focuses on the behavioral practices of women leaders using agentic strategies in their daily work. Women leaders with paradoxical thinking and strategic ability can strategically balance the tension between agency and community in interactions. Thus, it can more effectively coordinate the relationship between driving change and maintaining team harmony, and steadily promote performance improvement.

Paradoxical thinking (Zheng, Kark, et al., 2018) refers to female leaders combining proactiveness with amiability to counteract bias. The core cognitive framework of the model distinguishes between paradoxical thinking and dilemma thinking. Paradoxical thinking helps women leaders build psychological resilience, achieve compatibility between prejudice and leadership; The dilemma thinking may lead to the opposite result. (Zheng, Surgevil, et al., 2018) findings highlight how women leaders balance the paradoxical tension between agency and communion through creative integration strategies (such as soft delivery of hard messages). Women who exhibit both initiative and community are more likely to become leaders (Schock et al., 2019). Trzebiatowski et al. (2023) research shows that women leaders have achieved success in breaking through the glass ceiling by strategically applying gendered engagement strategies (skillfully tempered their agency with communion) to deal with double restraints.

In a competitive environment, women leaders may adopt an evolutionary disguise competitive strategy (showing competitiveness but not being tough). They need to find their place within the discourse system of punishment and expectation. This process is both a challenge and an opportunity for them to demonstrate their competitiveness (Mavin & Yusupova, 2024). Hierarchical competition may present a "downward" trend, where women leaders keep their distance from junior women. This creates "queen bee" phenomenon (Derks et al., 2016).

3.2 Literature Review: Core Debates and Evolution Trajectory

Through the above classification from the perspective of core theories, the key development context in this field has been clearly outlined, and its evolution shows a trend from simple judgment to multi-condition and dynamic evolution. The core debate revolves around the following three aspects.

3.2.1 Debate and Evolution I

The performance of women leaders has shifted from simple judgments (positive, negative, or no significant impact) to

situational dependence.

Earlier studies mainly explored the correlation between women leadership or the proportion in TMT on performance. The research results were mostly positive evaluations, with a few negative evaluations and conclusions of no significant impact. A meta-analysis of business cases involving female leaders found that the extent to which female leaders impact business performance varies (Hoobler et al., 2018). More complex contextual conditions, such as organizational gender equality culture, non-traditional projects, risky task, team coordination needs, paradoxical thinking, and strategic ability, jointly modulate the impact. In addition, methodological gap such as endogeneity problem (Sieweke et al., 2023) and subjective measure (Paustian-Underdahl et al., 2024) have drawn attention. Current research no longer simply judges the effects but systematically integrates multiple antecedents and emphasizes conditional dependencies (Buss et al., 2025).

3.2.2 Debate and Evolution II

The agentic and strategic abilities of women leaders have shifted from being constrained by stereotypes to integration and management.

Earlier studies focused more on gender stereotypes constraining the legitimacy and effectiveness of women leaders (Vial et al., 2016). However, a proactive trait of "self-reliance" has a positive impact on their evaluation (Schaumberg & Flynn, 2017). studies have shifted from the perspectives of paradoxical thinking theory (Zheng, Kark, et al., 2018), similarities between male and female leaders (Wille et al., 2018), and participation tactics (Trzebiatowski et al., 2023) to the agentic and contextualized strategic capabilities of women leaders. Women leaders have developed complex behavioral patterns that can effectively integrate and manage both "agency" and "warmth".

3.2.3 Debate and Evolution III

The source of women's leadership has deepened from "quantity" to "quality".

While "critical quantity" (roughly 30-33%) to acquire symbolic capital is important, research is gradually shifting towards emphasizing the crucial role of the "quality". The organizational, knowledge, and social capital of women directors are more important. Their political capital is particularly related to corporate social responsibility (Yang & Konrad, 2025). Furthermore, women directors build "bridges" between other networks of directors. It helps to access to key information and resources through social capital (Mateos De Cabo et al., 2022). It proves that quality is more explanatory than mere numerical statistics.

4. Discussion

4.1 Theoretical Gap

4.1.1 A Gap in the Theory of Multi-level Interaction.

Most of the existing research has focused on a single management level, while the influence of women leadership is a multi-level interactive process. There is a gap in the application of the relevant multi-level theory. For example, how do women CEOs or board members influence the career promotion and psychological safety of women middle managers by shaping organizational culture, formulating inclusive human resources policies, or serving as role models? How do middle women managers influence the creativity of project teams? How do changes in employees' effort affect organizational performance? While studies have explored the impact at the female founder and employee level (Kacperczyk et al., 2023), research on the broader cross-level impact within the organization (such as CEO, executive team, middle manager, and project team) is almost blank. Research has examined the cross-level effects of women's representation and diversity in senior management and hiring committees in the largest U.S. law firms (longitudinal data 2007-2015) (Dwivedi & Paolella, 2024), but there is still a lack of more systematic and updated cross-level research designs.

4.2 Methodological Gap

4.2.1 Limitations of Sample Time, Region and Institutional Cultural.

There is obvious sample bias in current studies, with the vast majority based on developed countries (Kirsch, 2018). There is still insufficient exploration of the Global South or different institutional and cultural contexts. Although the study of Colombia (Davila & Lluch, 2022) is an exception, cases from the Global South are still scarce. This severely limits the universality of the theory, as well as the understanding of institutional and cultural boundary conditions. There are significant

differences in the impact of gender quotas, institutional environments, ownership structures and women leadership on performance in state-owned enterprises, family businesses, or emerging industry. Therefore, the specific impact of national cultural values and gender diversity needs to be tested in a broader range of geographical and cultural samples.

In addition, the timeliness of some data is inadequate. For example, an analysis of start-ups in Portugal (longitudinal data from 2002 to 2012) fails to reflect recent trends and changes (Kacperczyk et al., 2023).

4.2.2 Endogeneity Issues and Causal Identification Challenges.

Existing studies generally face the "process black box" challenge caused by endogeneity issues, when exploring the women's leadership on performance. Business case researches often lack in-depth examination of the relative mechanisms. It is still unclear about how women leadership affects performance, which constitutes what is called the "missing variable" or "black box" problems (Hoobler et al., 2018). Existing studies often use cross-sectional or non-experimental designs to collect data, which are limited in inferring causal relationships and make it difficult to draw clear conclusions (Badura et al., 2018). A lot of the researches relies on correlational data. It will give rise to endogeneity problems. (Buss et al., 2025). Two panel studies, one with an analysis spanning 7 years (Baker et al., 2025) and the other involving 558 panel data analyses (Yang & Konrad, 2025), were both cautious about causal claims. Therefore, more reliable methods are needed to open the "black box" of process mechanisms.

4.3 Future Research

4.3.1 Capture the Multi-level Interactive Processes

This study is at the intersection of multiple disciplines and requires to capture the dynamic interaction processes at multiple levels (individuals, teams, organizations). Methodologically, it is necessary to explore how women senior leaders influence the innovative behavior of middle management teams and stimulate the creativity of employees by shaping organizational atmosphere or inclusive human resources practice policies. The samples need to be a multi-case comparative study, focusing on the longitudinal data, incorporating multiple variables such as age and appointment time, and tracking the transmission path of influence across organizational levels. Furthermore, it is necessary to verify the mechanism through which macro policies and micro behaviors affect performance, thereby uncovering the black box issues in the process.

4.3.2 Expand the Study in Global South Contexts

The review suggests that future research should focus on emerging economies such as China and delve deeply into the moderating effects of China's institutional and cultural context on women's leadership. For instance, under the " dual leadership of the Party and government " system implemented in state-owned enterprises, are the roles of women executives and their impact on performance particular? In order to reveal the moderating role of institutional context, cross-national comparative case studies could be designed to compare the effectiveness and influence of women directors' political capital in different institutional environments, such as China and the developed countries. It helps to fill the gaps in existing research in terms of geographical and institutional cultural context.

5. Conclusion

This systematic review points out that the impact of women's leadership on firm performance is not merely a simple positive or negative issue, but rather a complex process of dynamic regulation among bias challenges, resource complementarity, situational factors, and behavioral practices. Existing research has evolved from identifying role differences to exploring resource advantages and further to analyzing situational factors and behavioral practices. It also clearly pointed out the gaps that needed to be filled. Future research should follow this evolution and aim to construct more refined, process-oriented and context-sensitive theoretical models to illustrate under what conditions and through what means women leadership affect firm performance. The review indicates that future research should focus on multi-level interactive processes and enhance the exploration in the context of the Global South (developing country). It not only fills important theoretical and methodological gaps, but also provides practical insights for firms to more effectively stimulate the leadership potential of women.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Badura, K. L., Grijalva, E., Newman, D. A., Yan, T. T., & Jeon, G. (2018). Gender and leadership emergence: A meta-analysis and explanatory model. Personnel Psychology, 71(3), 335~367.

[2] Baker, M., Ali, M., Grabarski, M. K., & Konrad, A. (2025). Influence of Women in Leadership on Gender-Equality, Work–Life, and Domestic Violence Coping Practices: The Role of Chief Executive Officer Gender. Human Resource Management Journal, 35(4), 1032~1046.

[3] Buss, M., Andler, S., & Tiberius, V. (2025). Female leadership: An integrative review and research framework. The Leadership Quarterly, 36(3), 101858.

[4] Dai, Y., Byun, G., & Ding, F. (2019). The Direct and Indirect Impact of Gender Diversity in New Venture Teams on Innovation Performance. Entrepreneurship Theory and Practice, 43(3), 505~528.

[5] Dávila, C., & Lluch, A. (2022). Entrepreneurship in Emerging Markets: Female Entrepreneurs in Colombia since 1990. Business History Review, 96(2), 373~397.

[6] Derks, B., Van Laar, C., & Ellemers, N. (2016). The queen bee phenomenon: Why women leaders distance themselves from junior women. The Leadership Quarterly, 27(3), 456~469.

[7] Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. Strategic Management Journal, 33(9), 1072~1089.

[8] Dreher, G. F. (2003). Breaking the glass ceiling: The effects of sex ratios and work-life programs on female leadership at the top. Human Relations, 56(5), 541~562.

[9] Dwivedi, P., Misangyi, V. F., & Joshi, A. (2021). "Burnt by the spotlight": How leadership endorsements impact the longevity of female leaders. Journal of Applied Psychology, 106(12), 1885~1906.

[10] Dwivedi, P., & Paolella, L. (2024). Tick Off the Gender Diversity Box: Examining the Cross-Level Effects of Women's Representation in Senior Management. Academy of Management Journal, 67(4), 991~1023.

[11] Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. Psychological Review, 109(3), 573~598.

[12] Gamache, D. L., Devers, C. E., Klein, F. B., & Hannigan, T. (2023). Shifting perspectives: How scrutiny shapes the relationship between CEO gender and acquisition activity. Strategic Management Journal, 44(12), 3012~3041.

[13] Glass, C., & Cook, A. (2016). Leading at the top: Understanding women's challenges above the glass ceiling. The Leadership Quarterly, 27(1), 51~63.

[14] Glass, C., & Cook, A. (2018). Do women leaders promote positive change? Analyzing the effect of gender on business practices and diversity initiatives. Human Resource Management, 57(4), 823~837.

[15] Heilman, M. E. (2012). Gender stereotypes and workplace bias. Research in Organizational Behavior, 32, 113~135.

[16] Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. Journal of Applied Psychology, 92(1), 81~92.

[17] Hoobler, J. M., Masterson, C. R., Nkomo, S. M., & Michel, E. J. (2018). The Business Case for Women Leaders: Meta-Analysis, Research Critique, and Path Forward. Journal of Management, 44(6), 2473~2499.

[18] Jeong, S., Kang, S., & Byron, K. (2025). Bottom-up effects of female strategic leadership: Firm performance effects through employees. Strategic Management Journal, smj.70038.

[19] Joshi, M. P., & Diekman, A. B. (2022). My Fair Lady? Inferring Organizational Trust From the Mere Presence of Women in Leadership Roles. Personality and Social Psychology Bulletin, 48(8), 1220~1237.

[20] Kacperczyk, O., Younkin, P., & Rocha, V. (2023). Do Employees Work Less for Female Leaders? A Multi-Method Study of Entrepreneurial Firms. Organization Science, 34(3), 1111~1133.

[21] Kirsch, A. (2018). The gender composition of corporate boards: A review and research agenda. The Leadership Quarter-

ly, 29(2), 346~364.

[22] Knippen, J. M., Shen, W., & Zhu, Q. (2019). Limited progress? The effect of external pressure for board gender diversity on the increase of female directors. Strategic Management Journal, 40(7), 1123~1150.

[23] Mateos De Cabo, R., Grau, P., Gimeno, R., & Gabaldón, P. (2022). Shades of Power: Network Links with Gender Quotas and Corporate Governance Codes. British Journal of Management, 33(2), 703~723.

[24] Mavin, S., & Yusupova, M. (2024). Jostling Discourses of Competition: Women leaders self-positioning. Organization Studies, 45(6), 801~823.

[25] Ong, W. J. (2022). Gender-contingent effects of leadership on loneliness. Journal of Applied Psychology, 107(7), 1180~1202.

[26] Parker, O., Mui, R., & Titus, V. (2020). Unwelcome voices: The gender bias-mitigating potential of unconventionality. Strategic Management Journal, 41(4), 738~757.

[27] Paustian-Underdahl, S. C., Smith Sockbeson, C. E., Hall, A. V., & Halliday, C. S. (2024). Gender and evaluations of leadership behaviors: A meta-analytic review of 50 years of research. The Leadership Quarterly, 35(6), 101822.

[28] Post, C. (2015). When is female leadership an advantage? Coordination requirements, team cohesion, and team interaction norms: When is Female Leadership an Advantage? Journal of Organizational Behavior, 36(8), 1153~1175. https://doi.org/10.1002/job.2031

[29] Post, C., & Byron, K. (2015). Women on boards and firm financial performance: A meta-analysis. Academy of Management Journal, 58(5), 1546~1571.

[30] Post, C., Latu, I. M., & Belkin, L. Y. (2019). A Female Leadership Trust Advantage in Times of Crisis: Under What Conditions? Psychology of Women Quarterly, 43(2), 215~231.

[31] Powell, G. N., & Anthony Butterfield, D. (2015). The preference to work for a man or a woman: A matter of sex and gender? Journal of Vocational Behavior, 86, 28~37. https://doi.org/10.1016/j.jvb.2014.10.002

[32] Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. Research in Organizational Behavior, 28, 61~79.

[33] Ryan, M. K., & Haslam, S. A. (2005). The glass cliff: Evidence that women are over-represented in precarious leadership positions. British Journal of Management, 16(2), 81~90.

[34] Ryan, M. K., Haslam, S. A., Morgenroth, T., Rink, F., Stoker, J., & Peters, K. (2016). Getting on top of the glass cliff: Reviewing a decade of evidence, explanations, and impact. The Leadership Quarterly, 27(3), 446~455.

[35] Sánchez, C. M., & Lehnert, K. (2019). The unbearable heaviness of leadership: The effects of competency, negatives, and experience on women's aspirations to leadership. Journal of Business Research, 95, 182~194.

[36] Schaumberg, R. L., & Flynn, F. J. (2017). Self-reliance: A Gender Perspective on its Relationship to Communality and Leadership Evaluations. Academy of Management Journal, 60(5), 1859~1881.

[37] Schock, A.-K., Gruber, F. M., Scherndl, T., & Ortner, T. M. (2019). Tempering agency with communion increases women's leadership emergence in all-women groups: Evidence for role congruity theory in a field setting. The Leadership Quarterly, 30(2), 189~198.

[38] Schultheiss, D. E. (2021). Shining the light on women's work, this time brighter: Let's start at the top. Journal of Vocational Behavior, 126, 103558.

[39] Sergent, K., & Stajkovic, A. D. (2020). Women's leadership is associated with fewer deaths during the COVID-19 crisis: Quantitative and qualitative analyses of United States governors. Journal of Applied Psychology, 105(8), 771~783.

[40] Sidhu, J. S., Feng, Y., Volberda, H. W., & Van Den Bosch, F. A. J. (2021). In the Shadow of Social Stereotypes: Gender diversity on corporate boards, board chair's gender and strategic change. Organization Studies, 42(11), 1677~1698.

[41] Sieweke, J., Bostandzic, D., & Smolinski, S.-M. (2023). The influence of top management team gender diversity on firm performance during stable periods and economic crises: An instrumental variable analysis. The Leadership Quarterly, 34(5), 101703.

[42] Stainback, K., Roberts, H., & Biswas, P. K. (2024). Women's representation in managerial hierarchies: An examination

of trickle-down and pipeline effects. Human Resource Management, 63(4), 619~637.

[43] Terpstra-Tong, J. L. Y., Treviño, L. J., Yaman, A. C., Froese, F. J., Ralston, D. A., Bozionelos, N., Furrer, O., Tjemkes, B., León-Darder, F., Li, Y., Fu, P., Molteni, M., Palmer, I., Tučková, Z., Szabo, E., Poeschl, G., Hemmert, M., De La Garza Carranza, M. T., Suzuki, S., … Baltazar Herrera, M. (2025). Gender composition at work and women's career satisfaction: An international study of 35 societies. Human Resource Management Journal, 35(2), 397~427.

[44] Trzebiatowski, T., McCluney, C., & Hernandez, M. (2023). Managing the Double Bind: Women Directors' Participation Tactics in the Gendered Boardroom. Organization Science, 34(2), 801~827.

[45] Vial, A. C., Napier, J. L., & Brescoll, V. L. (2016). A bed of thorns: Female leaders and the self-reinforcing cycle of illegitimacy. The Leadership Quarterly, 27(3), 400~414.

[46] Wille, B., Wiernik, B. M., Vergauwe, J., Vrijdags, A., & Trbovic, N. (2018). Personality characteristics of male and female executives: Distinct pathways to success? Journal of Vocational Behavior, 106, 220~235.

[47] Wu, J., Richard, O. C., Triana, M. D. C., & Zhang, X. (2022). The performance impact of gender diversity in the top management team and board of directors: A multiteam systems approach. Human Resource Management, 61(2), 157~180.

[48] Yang, Y., & Konrad, A. M. (2025). Effects of women on corporate boards: An integrative review from a political capital perspective. The Leadership Quarterly, 36(1), 101841.

[49] Zheng, W., Kark, R., & Meister, A. L. (2018). Paradox versus dilemma mindset: A theory of how women leaders navigate the tensions between agency and communion. The Leadership Quarterly, 29(5), 584~596.

[50] Zheng, W., Surgevil, O., & Kark, R. (2018). Dancing on the Razor's Edge: How Top-Level Women Leaders Manage the Paradoxical Tensions between Agency and Communion. Sex Roles, 79(11~12), 633~650.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

A Critical Observation of SpaceX's Founding Phase

Ziyao Yang*

University of Chinese Academy of Social Sciences, Chaoyang District, Beijing, 100102, China

Corresponding author: Ziyao Yang, ziy1.@stir.ac.uk

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Since the 21st century, the global aerospace industry has transitioned from state monopoly to public-private collaboration, giving rise to the "New Space Economy." SpaceX, founded in 2002, emerged as a pivotal disruptor, evolving from a struggling startup to the world's most valuable private aerospace firm. This study conducts a critical single-case analysis of SpaceX's formative decade (2002–2012), spanning its founding to the Dragon spacecraft's historic ISS docking. Using process tracing, systematic literature review, and critical discourse analysis, the paper examines SpaceX's entrepreneurial origins, growth strategy, financing models, digital capability building, and organizational evolution. The findings reveal that while SpaceX's success stemmed from iterative innovation, vertical integration, milestone-based government financing, and mission-driven culture, it was highly contingent on Elon Musk's personal wealth, favorable geopolitical timing, and extreme risk tolerance. Critical limitations include overreliance on government contracts, autocratic leadership vulnerabilities, and unsustainable work intensity. Drawing on these insights, the study provides actionable recommendations for Chinese hard tech startups, emphasizing the need to balance long-term vision with financial discipline, build strategic partnerships, foster inclusive organizational cultures, and proactively engage with regulators. This research contributes to a nuanced understanding of hard tech entrepreneurship, highlighting both replicable best practices and context-specific constraints of the SpaceX model.

Keywords: SpaceX; Hard Tech Entrepreneurship; Startup Strategy; Public-Private Partnership; Iterative Innovation; Entrepreneurial Financing

Published: Jun 17, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1509

1. Introduction

1.1 Research Background and Significance

Since the 21st century, the global aerospace industry has shifted from state monopoly to public-private collaboration, driven by technology convergence and ecosystem integration, leading to the "New Space Economy" and over 90% reduction in space access costs. SpaceX, founded in 2002, is a key disruptor, evolving from a struggling startup to the world's most valuable private aerospace company, illustrating hard tech entrepreneurship.

Amid China's push for technological self-reliance, hard tech startups in semiconductors, AI, biomedicine, and aerospace face barriers like protracted R&D cycles, high capital needs, and innovation-market misalignment. While existing research documents SpaceX's successes, few studies critically examine its model's limitations and generalizability to non-US contexts. This study addresses this gap by analyzing SpaceX's early strategies to identify replicable best practices and context-specific advantages, offering insights for Chinese hard tech startups in long-cycle growth.

1.2 Rationale for Selecting the First 10 Years of SpaceX's Start-up Period

This study focuses on SpaceX's formative decade (2002-2012), from its founding to the Dragon's historic ISS docking. This period represents the complete "0-to-1" lifecycle of a hard tech startup, marked by launch failures, near-bankruptcy, and the NASA CRS contract. Three justifications support this timeframe: it captures opportunity identification, technological iteration, and market validation; strategic decisions laid the foundation for SpaceX's dominance; and post-2012, SpaceX entered a scaling phase outside early-stage entrepreneurship research.

1.3 Research Methods and Data Sources

This paper uses a single-case study method, integrating process tracing, systematic literature review, and critical discourse analysis. Data is triangulated from three sources: 10 peer-reviewed academic articles on SpaceX and aerospace entrepreneurship indexed in the Web of Science Core Collection, primary documents (e.g., SpaceX and NASA reports), and authoritative media. This multi-source verification ensures reliability and validity, while the critical framework allows a nuanced assessment of SpaceX's model's strengths and weaknesses.

2. Entrepreneurial Origin and Development Process

2.1 Identification and Origin of Innovation and Entrepreneurship Opportunities

SpaceX's genesis traces to Elon Musk's 2001 "Mars Oasis" project, which aimed to send a mini-greenhouse to Mars to reignite public interest in space exploration. While attempting to purchase refurbished Russian intercontinental ballistic missiles (ICBMs) for the mission, Musk discovered a critical market inefficiency: traditional aerospace contractors charged exorbitant prices despite rockets' manufacturing costs accounting for 2% of total launch expenses. The root cause, he concluded, was the industry's reliance on disposable rocket architecture.

This insight revealed a disruptive entrepreneurial opportunity: developing reusable rockets to cut launch costs by an order of magnitude, thereby democratizing access to space. Unlike traditional aerospace firms focused on incremental government contracts, Musk anchored SpaceX's mission to "making humanity a multi-planetary species." This transcendent vision served as a powerful magnet for top talent and patient capital, differentiating SpaceX from conventional startups (Muegge & Reid, 2019).

Critical Analysis: While Musk's opportunity identification was prescient, it was highly contingent on his unique position as a wealthy serial entrepreneur with access to networks and capital unavailable to most founders. His ability to personally absorb $140 million in losses during the early years created a buffer that no ordinary hard tech startup could replicate.

2.2 Development Process in the Founding Period

SpaceX was founded in June 2002 with Musk investing $100 million from his PayPal proceeds. Initially operating from a warehouse, its first rocket was the Falcon 1. The company faced setbacks, with three failed launches from 2006 to 2008, draining funds and forcing Musk to inject another $40 million. The fourth Falcon 1 launch in September 2008 succeeded, becoming the first private liquid-fuel rocket to reach orbit. This attracted NASA, which awarded a $1.6 billion contract for ISS resupply in December 2008. Later milestones included the Falcon 9 debut in 2010 and Dragon-ISS docking in 2012, solidifying SpaceX's role in aerospace.

Critical Analysis: SpaceX's survival was not solely due to technological prowess but also to extraordinary luck and timing. The fourth Falcon 1 launch succeeded with only hours of funding remaining; a single additional failure would have resulted in bankruptcy. Furthermore, the NASA CRS contract was awarded during a period when the US government was seeking to reduce its dependence on Russian Soyuz spacecraft following the retirement of the Space Shuttle program, creating a unique geopolitical window of opportunity.

2.3 Evaluation of the Founder and Team Capabilities

SpaceX's survival and success are inextricably linked to Musk's entrepreneurial leadership. As a serial entrepreneur with proven expertise in scaling technology ventures, he combined exceptional technical intuition with ruthless execution. His ability to decompose audacious long-term goals into incremental, measurable objectives kept the team focused amid repeated failures. Equally critical was his extraordinary resilience: during the 2008 crisis, he simultaneously navigated SpaceX's near-collapse and Tesla's financial troubles, demonstrating unparalleled pressure tolerance.

Musk also assembled a world-class core team, poaching top engineers from Boeing, Lockheed Martin, and NASA. These individuals shared his vision and embraced a high-risk, high-reward work culture. SpaceX's flat organizational structure—with minimal management layers and direct communication channels between engineers and executives—enabled rapid decision-making and agile problem-solving, a stark contrast to the bureaucratic hierarchies of traditional aerospace incumbents.

Critical Analysis: While Musk's leadership was instrumental to SpaceX's success, it also created significant vulnerabilities. His autocratic decision-making style, while efficient in crisis situations, has led to costly strategic missteps. Furthermore, the company's overreliance on Musk's personal brand and leadership creates a substantial succession risk, with no clear plan for leadership continuity.

3. Enterprise Growth Strategy Analysis

3.1 Strategic Disassembly Based on Ganz's Four-Element Model

Ganz's Four-Element Strategic Model posits that a firm's competitive advantage derives from the alignment of four interdependent choices: customer, technology, competition, and organization. This framework provides a systematic lens to analyze SpaceX's early-stage strategy.

Customer Choice: SpaceX adopted a dual-track "government + commercial" customer strategy. In its nascent stage, when the commercial launch market was underdeveloped, government contracts provided stable cash flow and credibility. NASA's CRS contract was transformative, not only delivering $1.6 billion in revenue but also validating SpaceX's technology in a high-stakes regulatory environment. Concurrently, SpaceX targeted underserved commercial segments, particularly small satellite operators ignored by incumbents. This hybrid model balanced short-term survival with long-term market expansion, with commercial customers gradually accounting for a growing share of revenue as the industry matured.

Technology Choice: SpaceX's technology strategy centered on "cost leadership through reusability." Rejecting the industry's obsession with cutting-edge but expensive aerospace-grade components, it prioritized modular design, commercial off-the-shelf (COTS) parts, and vertical integration. By manufacturing 80% of rocket components in-house, SpaceX reduced supply chain costs and accelerated iteration cycles. The cornerstone of its technology roadmap was reusable rocket technology, which promised to slash launch costs by 90% by recovering and refurbishing first-stage boosters (Zimcik, 2017). This focus on cost reduction rather than absolute performance was the defining disruptive feature of SpaceX's innovation.

Competition Choice: SpaceX pursued a differentiated low-cost strategy to disrupt the entrenched monopoly of Boeing and Lockheed Martin's United Launch Alliance (ULA). Initially, it targeted the small satellite launch market, a niche overlooked by incumbents focused on large government payloads. By demonstrating reliable, low-cost launches with Falcon 1, SpaceX built brand equity and scaled up to the medium-lift market with Falcon 9. Its ability to undercut ULA's prices by 50% while offering higher launch frequencies allowed it to rapidly capture market share, breaking the decades-old government contractor monopoly.

Organizational Choice: SpaceX's organizational design was optimized for speed and innovation. Its flat structure eliminated bureaucratic red tape, enabling engineers to escalate issues directly to Musk. The company fostered a mission-driven culture that celebrated risk-taking and normalized failure as a learning opportunity. Generous equity incentives aligned employee interests with long-term company success, attracting talent willing to trade job security for the chance to revolutionize space exploration.

Critical Analysis of Ganz's Four-Element Strategy: SpaceX's strategic alignment came with significant trade-offs and context dependencies. First, its customer strategy was overly reliant on government largesse, with over 70% of 2012 revenue coming from NASA contracts, creating vulnerability to policy shifts. Second, the COTS parts approach initially caused reliability issues and raised safety concerns as the company entered human spaceflight. Third, its competitive advantage was as much a product of ULA's monopoly-induced inefficiency as its own innovation—a condition not present in more competitive industries. Finally, its high-intensity, mission-driven culture led to extreme work demands, high turnover, and a lack of diversity, which may hinder long-term innovation.

3.2 Technological Innovation and Business Growth of Startups

SpaceX's innovation followed an iterative "test-fail-learn" cycle, contrasting with the traditional aerospace industry's "design-build-test-fly" approach. Rooted in Silicon Valley's agile principles, this methodology compressed R&D timelines and reduced costs. For example, the Merlin engine underwent over 100 design iterations in five years, with each test informing improvements.

This iterative innovation drove business growth. The 2008 Falcon 1 success validated SpaceX's capabilities and secured the NASA CRS contract. The 2010 Falcon 9 debut expanded its market to medium and heavy payloads, enabling competition in commercial satellite launches. By 2012, SpaceX had over 40 launch contracts worth $4 billion, with revenue rising from zero to over $200 million and employees growing from 100 to 1,800.

Crucially, SpaceX embedded failure into its innovation process. As Vittori et al. (2024) note, systematically analyzing failures and iterating rapidly was a key competitive advantage, with each failure leading to design improvements that enhanced reliability and cut costs over time.

Critical Analysis: While the "test-fail-learn" model worked well for SpaceX, it is not universally applicable to all hard tech industries. In sectors such as nuclear energy or pharmaceuticals, where failure can have catastrophic consequences, a more cautious approach to innovation is necessary. Furthermore, the model requires significant financial resources to absorb the costs of repeated failures, which most startups do not have.

3.3 Development Potential

By 2012, SpaceX had laid the groundwork for exponential growth. Its reusable rocket technology roadmap promised to further disrupt the launch market, opening up new opportunities in satellite internet, space tourism, and deep space exploration. The success of the Dragon spacecraft positioned SpaceX to compete for NASA's Commercial Crew Program, which aimed to restore U.S. human spaceflight capability. Additionally, its vertical integration model and manufacturing expertise gave it a sustainable cost advantage over competitors.

4. Entrepreneurial Finance and Financing Strategy Analysis

4.1 Financial Situation in the Start-up Period

SpaceX's early financial history was characterized by extreme volatility and existential risk. From 2002 to 2008, the company generated no revenue and relied entirely on Musk's personal investment. By the third Falcon 1 failure in August 2008, Musk had poured $140 million into SpaceX, nearly depleting his entire net worth. The company was days away from bankruptcy when the fourth Falcon 1 launch succeeded, a turnaround that has become legendary in entrepreneurial circles.

The 2008 NASA CRS contract marked a turning point in SpaceX's financial fortunes. The $1.6 billion contract, paid incrementally upon mission completion, provided a stable revenue stream that allowed the company to scale operations and invest in next-generation technology. In 2010, SpaceX raised its first external funding round of $20 million from Founders Fund, valuing the company at $275 million. Subsequent rounds in 2011 and 2012 brought total external funding to $170 million, with investors including Draper Fisher Jurvetson and Google Ventures.

Despite these infusions, SpaceX remained unprofitable throughout its first decade. The company reinvested nearly all revenue into R&D and manufacturing capacity expansion, prioritizing long-term growth over short-term profitability. This "burn rate" strategy was typical of hard tech startups, where significant upfront investment is required to overcome technical barriers and achieve economies of scale.

4.2 Analysis of Financing Process and Financing Method Choice

SpaceX's financing strategy evolved in three distinct phases, reflecting the changing risk profile of the company as it progressed from technical validation to commercialization.

Founder Self-Financing Phase (2002–2008): In the pre-revenue technical validation stage, SpaceX was unable to attract external venture capital, which typically avoids high-risk, long-cycle hard tech projects. Musk's willingness to invest his entire fortune was critical to the company's survival. This approach gave him complete control over strategic decisions, allowing him to prioritize long-term vision over investor demands for quick returns. However, it also exposed the company to existential risk, as a single additional failure would have resulted in bankruptcy.

Government Contract Financing Phase (2008–2010): NASA's CRS contract represented a novel form of "milestone-

based" financing for hard tech startups. Unlike traditional government grants, which are awarded upfront with minimal accountability, the CRS contract tied payments to specific technical and operational milestones. This model de-risked the investment for both SpaceX and the government: NASA only paid for successful missions, while SpaceX gained access to capital without diluting equity. As King et al. (2024) argue, this public-private partnership model has become a blueprint for stimulating innovation in capital-intensive industries.

Venture Capital Financing Phase (2010–2012): With the technical success of Falcon 1 and Falcon 9 and the security of the NASA contract, SpaceX became an attractive investment for venture capital firms. The 2010 Founders Fund round marked the beginning of a series of increasingly large funding rounds that would eventually value the company at over $100 billion. Venture capital provided the capital needed to scale manufacturing, accelerate R&D on reusable rockets, and expand into new markets. In addition to capital, investors brought valuable industry connections and management expertise, helping SpaceX transition from a startup to a mature enterprise.

Beyond these traditional methods, SpaceX explored innovative financing mechanisms for its long-term Mars colonization goals. Harris and Wonglimpiyarat (2023) highlight the potential of cryptocurrency, initial coin offerings (ICOs), and crowdfunding as alternative funding sources for deep space missions. While SpaceX did not adopt these methods in its early years, they represent promising avenues for future hard tech ventures pursuing transformative, capital-intensive projects.

Critical Analysis: SpaceX's financial strategy was too risky for most startups. Its survival without revenue for six years relied on Musk's personal wealth, a rare advantage. The "burn rate" model demands continuous external financing, risking founder control dilution and a short-term growth focus. While the milestone-based financing model is promising, it has limitations. Government contracts are competitive and politically influenced, making them unreliable for startups. The model also requires government technical expertise to set and evaluate milestones, which is often lacking. In countries with underdeveloped procurement systems, it may be unfeasible.

This capital-driven innovation model has inherent drawbacks: first, it is extremely difficult to replicate, as it depends on a unique combination of a charismatic founder, favorable market conditions, and patient capital; second, if a company cannot achieve sustainable profitability eventually, it will face liquidation by capital; third, as the company scales, the founder's vision often conflicts with investors' short-term profit motives; and fourth, it requires investors to have deep technical understanding of frontier technologies, a capability that is still rare in most markets.

5. Digital Capability and Enterprise Organizational Growth

5.1 Construction of Digital Capability

SpaceX's digital transformation was a cornerstone of its ability to innovate rapidly and reduce costs. The company integrated digital technologies across the entire product lifecycle, from design and manufacturing to testing and launch operations.

In the design phase, SpaceX employed advanced computer-aided design (CAD) and computer-aided engineering (CAE) tools to create fully digital prototypes of rockets and spacecraft. Engineers could simulate thousands of launch scenarios, identify potential failures, and optimize designs before physical prototypes were built. This digital-first approach cut design cycles by 50% and reduced the number of costly physical tests. A centralized product data management (PDM) system ensured that all teams worked from the same up-to-date design files, eliminating errors and improving collaboration.

In manufacturing, SpaceX was an early adopter of additive manufacturing (3D printing) technology. 3D printing allowed the company to produce complex, high-performance components that would be impossible or prohibitively expensive to manufacture using traditional methods. For example, the Merlin engine's combustion chamber was 3D printed as a single piece, reducing the number of parts from over 100 to 1 and cutting manufacturing time from months to days. SpaceX also implemented digital manufacturing execution systems (MES) to automate production lines, track inventory in real time, and improve quality control.

In testing and launch operations, SpaceX leveraged big data and artificial intelligence (AI) to monitor rocket performance. Thousands of sensors on each rocket generated terabytes of data during flight, which was analyzed in real time to detect anomalies and inform design improvements. AI algorithms were used to optimize rocket landing trajectories, significantly increasing the success rate of first-stage booster recoveries. This data-driven approach to engineering allowed SpaceX to

continuously improve its products at a pace unmatched by traditional aerospace firms.

Critical Analysis: While SpaceX's digital capabilities are impressive, they also create new vulnerabilities. The company's reliance on digital systems makes it susceptible to cyberattacks, which could have catastrophic consequences for launch operations.

5.2 Organizational Growth

SpaceX evolved from a linear to a matrix structure as it grew, while keeping its flat, agile culture. Initially, with under 100 employees, Musk managed directly for fast, bureaucracy-free decisions. By 2012, with over 1,800 staff, this became unsustainable, leading to a matrix organization with functional departments and cross-functional project teams. This balanced expertise and focus, enabling efficient management of multiple programs. SpaceX maintained flat management, with only three to four levels between engineers and Musk, and weekly all-hands meetings encouraged open communication and direct challenges.

SpaceX's organizational culture was defined by its mission-driven focus, bias for action, and tolerance for failure. Employees were motivated by the opportunity to work on transformative projects rather than just financial compensation. The company celebrated risk-taking and viewed failures as valuable learning experiences, a mindset that was essential for innovation in a high-risk industry. This culture attracted top talent and fostered a high-performance environment where employees were willing to work long hours and take on significant responsibility.

Critical Analysis: The matrix structure, while effective for managing complex projects, can also lead to confusion and conflict over roles and responsibilities. Furthermore, SpaceX's flat management style, while promoting agility, can result in a lack of clear career paths for employees, contributing to high turnover. The company's culture of extreme work intensity has also been linked to mental health issues among employees, raising questions about the long-term sustainability of this model.

5.3 Founder's Growth

SpaceX's success was also a story of Elon Musk's personal growth as an entrepreneur and leader. When he founded SpaceX, Musk was a successful internet entrepreneur with no experience in aerospace engineering or large-scale manufacturing. Over the course of the first decade, he transformed himself into a world-class aerospace executive, mastering the technical complexities of rocket science and learning to manage a rapidly growing organization.

Musk's leadership style evolved from hands-on micromanagement to a more strategic approach. In the early years, he was deeply involved in every aspect of rocket design and testing, often working alongside engineers on the factory floor. As the company grew, he learned to delegate operational responsibilities to trusted lieutenants, focusing his attention on long-term strategy, vision, and key technical decisions. He also developed expertise in government relations, navigating the complex regulatory landscape of the aerospace industry and building strong partnerships with NASA and other government agencies.

Critical Analysis: While Musk's personal growth was impressive, his leadership style remains highly controversial. His tendency to make impulsive decisions and set unrealistic deadlines has led to significant delays and cost overruns on several projects. Furthermore, his public statements on social media have sometimes created legal and reputational risks for the company, highlighting the dangers of overreliance on a single charismatic leader.

6. Countermeasures for the Startups and Suggestions

6.1 Challenges Faced by Enterprises in the Start-up Period

SpaceX's early struggles highlight the universal challenges facing hard tech startups. First and foremost is technical risk: hard tech innovation involves pushing the boundaries of science and engineering, and failure is inevitable. SpaceX's three launch failures are a testament to the difficulty of developing new technologies in complex, high-stakes industries. Second is financial risk: hard tech startups require massive upfront capital investment, and revenue generation is often delayed by years. Many promising ventures fail not because of flawed technology, but because they run out of money before achieving commercial viability.

Third is talent acquisition: hard tech startups compete with large, established companies and academic institutions for top technical talent. Attracting and retaining skilled engineers and scientists is particularly challenging for early-stage startups, which often cannot match the salaries and benefits offered by incumbents. Fourth is market access: hard tech products often

require regulatory approval and face high barriers to entry, especially in industries like aerospace, healthcare, and energy. Building relationships with customers, regulators, and industry partners is critical for success.

Fifth is the inherent tension between founder vision and capital interests. As startups scale, investors often push for short-term profits at the expense of long-term transformative goals. This conflict can derail the original mission of the company, especially in frontier tech sectors where breakthroughs require decades of sustained investment.

6.2 Development Suggestions

Drawing on SpaceX's experience and its limitations, this paper offers five key recommendations for hard tech startups seeking sustainable growth. First, adopt a long-term vision but maintain financial discipline. While a clear, inspiring mission is essential for motivating employees and attracting investors, startups must also manage their cash flow carefully and avoid overreliance on a single source of funding. SpaceX's near-bankruptcy in 2008 serves as a cautionary tale about the dangers of excessive financial risk-taking.

Second, focus on core technology but build strategic partnerships. Instead of trying to do everything in-house, startups should identify their core competencies and partner with other companies or research institutions for non-core activities. This can reduce costs, accelerate development, and mitigate technical risks. While SpaceX's vertical integration strategy was successful, it required significant capital and expertise that most startups do not have.

Third, build a diverse and inclusive organizational culture. While mission-driven cultures are effective for driving innovation, they must also prioritize employee well-being and work-life balance. A diverse workforce brings different perspectives and experiences, leading to better decision-making and innovation. Startups should also establish clear career paths and professional development opportunities to retain top talent.

Fourth, engage proactively with regulators and policymakers. Hard tech industries are often heavily regulated, and startups that fail to understand and comply with regulations can face significant delays and costs. By engaging with regulators early in the development process, startups can help shape regulations that are supportive of innovation while ensuring public safety.

Fifth, strategically balance founder vision with investor expectations. Founders should seek out patient capital from investors who share their long-term vision, rather than those focused solely on quick returns. They should also structure financing rounds to preserve voting control and protect the company's core mission. Additionally, startups should proactively identify and demonstrate real-world use cases for their technology to validate their value proposition to both investors and the public, as SpaceX did with Starlink.

6.3 Two Development Paths and Countermeasures

After successfully navigating the start-up phase, hard tech startups typically face two strategic options: independent growth or acquisition. The independent growth path is suitable for companies with proprietary technology, large addressable markets, and access to sufficient capital. SpaceX chose this path, leveraging its technological advantages to become a dominant player in the global aerospace industry. Startups pursuing independent growth should focus on scaling operations, expanding into new markets, and building sustainable competitive advantages. They should also invest in building strong management teams and corporate governance structures to reduce reliance on the founder.

The acquisition path is a viable alternative for startups with valuable technology but limited access to capital or market reach. Being acquired by a larger company can provide access to resources, distribution channels, and customer bases that would be difficult to build independently. For example, many semiconductor and biotech startups are acquired by larger firms that can commercialize their technology more effectively. Startups considering acquisition should focus on developing clear, defensible intellectual property and building relationships with potential acquirers early on. They should also negotiate favorable terms that preserve the autonomy of the founding team and ensure the continued development of the technology.

7. Enlightenment

SpaceX's first decade of innovation and entrepreneurship offers both valuable lessons and important caveats for Chinese hard tech startups operating in an increasingly competitive global landscape. On the one hand, SpaceX demonstrates that long-term vision, iterative innovation, and public-private partnerships can drive transformative change in even the most entrenched industries. Its success shows that hard tech startups can disrupt incumbents by redefining the problem and focusing on

customer needs rather than incremental improvements.

SpaceX's breakthrough represents a paradigm shift in how frontier innovation is funded and executed. For the first time, private capital has successfully led a megaproject that was previously the exclusive domain of nation-states. This has not only accelerated the development of the aerospace industry and spawned a whole ecosystem of related businesses but has also deepened the integration between capital and technology to an unprecedented degree. It has proven that with the right model, social capital can be mobilized to tackle humanity's grand challenges, breaking the ideological shackles that certain industries must be controlled by the state.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Jenaro, G., Hidalgo, A. (2026). 'How technology convergence and entrepreneurial leadership disrupted the launch industry: the case of SpaceX'. International Entrepreneurship and Management Journal 22(2), Article 74. doi: http://dx.doi.org/10.1007/s11365-026-01198-6

[2] Muegge, S., Reid, E. (2019). 'Elon Musk and SpaceX: A Case Study of Entrepreneuring as Emancipation'. Technology Innovation Management Review 9(8), pp. 18-29. doi: http://dx.doi.org/10.22215/timreview/1296

[3] Zimcik, P. (2017). 'GROWTH OF A NEW MARKET: INNOVATION IN SPACE INDUSTRY'. 2nd Annual International Scientific Conference of Business Economics, Management and Marketing (ISCOBEMM), Zajeci, CZECH REPUBLIC, MAY 25-26, 2017, pp. 328-338.

[4] Lichtenthaler, U. (2018). 'The world's most innovative companies: a meta-ranking'. Journal of Strategy and Management 11(4), pp. 497-511. doi: http://dx.doi.org/10.1108/JSMA-04-2018-0041

[5] Vittori, D., Natalicchio, A., Panniello, U., Petruzzelli, A., Albino, V. and Cupertino, F. (2024). 'Failure is an option: How failure can lead to disruptive innovations'. Technovation 129, Article 102898. doi: http://dx.doi.org/10.1016/j.technovation.2024.102898

[6] King, R., Smith, D. and Williams, G. (2024). 'The final frontier? Mechanisms to stimulate investment in the commercial space travel industry'. Accounting Auditing & Accountability Journal 37(5), pp. 1403-1427. doi: http://dx.doi.org/10.1108/AAAJ-09-2023-1611

[7] Harris, W., Wonglimpiyarat, J. (2023). 'Aerospace financing towards Mars commercialisation'. Foresight 25(4), pp. 617-630. doi: http://dx.doi.org/10.1108/FS-09-2022-0131

[8] King, R., Smith, D. and Williams, G. (2024). 'The final frontier? Mechanisms to stimulate investment in the commercial space travel industry'. Accounting Auditing & Accountability Journal 37(5), pp. 1403-1427. doi: http://dx.doi.org/10.1108/AAAJ-09-2023-1611

[9] Harris, W., Wonglimpiyarat, J. (2023). 'Aerospace financing towards Mars commercialisation'. Foresight 25(4), pp. 617-630. doi: http://dx.doi.org/10.1108/FS-09-2022-0131

[10] Steers, R., Nardon, L. and Sanchez-Runde, C. (2016). 'Managing work and motivation'. Management Across Cultures: Developing Global Competencies, 3rd Edition, pp. 262-291. ISBN 978-1-107-15079-9; 978-1-316-60403-8.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Mapping Legitimacy Research in Business and Management: A CiteSpace-Based Bibliometric Review of Conceptual Diffusion and Thematic Evolution

Zerong Huang[1], Doudou Zi[1], Tao Wang[1,2]*

1. School of Innovation and Entrepreneurship, Guangzhou Xinhua University, Guangzhou 510520, Guangdong, China
2. Faculty of Business and Communications, INTI International University, Nilai 71800, Negeri Sembilan, Malaysia

Corresponding author: Tao Wang, elontao020@gmail.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: This study examines the diffusion and thematic evolution of legitimacy research in business and management. Based on 6,034 records from the Web of Science Core Collection covering 1997-2025, the paper uses descriptive statistics, co-citation analysis, keyword clustering, burst detection, and thematic evolution mapping to identify the field's publication patterns, knowledge structure, and research frontiers. The findings show that legitimacy has expanded from its institutional-theory foundations into entrepreneurship, CSR, sustainability reporting, ESG, governance accountability, international business, and digital platform studies. Rather than replacing earlier traditions, newer themes reconnect with classic questions of organizational recognition, audience judgment, disclosure, responsibility, and resource mobilization. The review further suggests that CSR, carbon disclosure, ESG performance, green innovation, artificial intelligence, and gender diversity have become active frontier topics. The paper argues that future research should specify the type of legitimacy being studied, the audiences making legitimacy judgments, and the evidence through which legitimacy is acquired, maintained, challenged, repaired, or lost.

Keywords: Legitimacy; Bibliometric Analysis; CiteSpace; Business and Management; Conceptual Diffusion; Thematic Evolution
Published: Jun 17, 2026
DOI: https://doi.org/10.62177/apemr.v3i3.1508

1. Introduction

Legitimacy is a key concept in institutional theory, but its application has gradually expanded beyond the scope of early institutionalist research. Early institutionalist studies used this concept to explain that an organization's survival depends not only on efficiency but also on whether it is perceived as legitimate, acceptable, and meaningful within a specific social order [1–4]. This perspective remains relevant today, though the contexts in which legitimacy plays a role have become more complex. Today, firms must not only secure the approval of regulators and investors but also respond to the expectations of a diverse array of stakeholders, including consumers, communities, the media, rating agencies, platform users, and activist groups. Consequently, legitimacy is no longer merely a theoretical term; it has become a practical prerequisite for organizational survival, market access, and resource mobilization.

The term "legitimacy" now carries too many meanings simultaneously. In some studies, legitimacy refers to compliance with institutional rules; in others, it is almost equated with stakeholder endorsement, moral acceptance, public trust, reputation, or

credibility. While these concepts clearly overlap, the widespread dissemination and use of the concept of legitimacy may also cause it to lose its analytical power. This paper takes this critical point as its starting point: legitimacy is influential precisely because it bridges numerous discussions within management studies; however, it is this very versatility that may undermine the precision it originally possessed as an analytical concept [13–19]. This characteristic is evident across multiple fields. Entrepreneurship research employs legitimacy to explain how startups are perceived by investors and customers; corporate social responsibility and sustainability research utilize legitimacy to analyze information disclosure, ESG ratings, and greenwashing. International business researchers leverage legitimacy to examine host country acceptance and institutional complexity. While they share the same conceptual vocabulary, they often address different questions.

Existing conceptual reviews have clarified many foundational issues in the field. They explain how organizations acquire, maintain, or lose legitimacy, and they clarify why legitimacy differs from reputation, status, and trust [13–15]. However, such reviews do not readily capture the overall structure of the field as it has emerged over the past three decades at the level of journals, citations, and keywords. For this reason, bibliometric maps hold significant value. Maps and networks do not automatically provide explanations, but they can reveal where the research community congregates, where it converges, and where the concept of legitimacy is being continually expanded.

Therefore, this paper examines research related to legitimacy in the fields of business and management from 1997 to 2025. The paper addresses three central questions: First, which knowledge traditions constitute the primary organizational foundations of the field; second, how clusters of major research themes and keywords have evolved; and third, how emerging fields such as corporate social responsibility, ESG, sustainability reporting, and governance accountability have reshaped the use of the concept of legitimacy. The aim of this paper is to reveal how the concept of legitimacy has diffused and been reconfigured, and to identify areas where future research requires greater theoretical awareness and conceptual rigor [29–34].

2. Research Overview and Theoretical Background

2.1 The Institutional Context of the Concept of Legitimacy

Institutional theory provides a key framework for understanding legitimacy. Meyer and Rowan note that an organization's formal structure is not merely intended to enhance efficiency; it sometimes functions as a socially accepted "rationalization myth" [1]. DiMaggio and Powell further explain why organizations within the same institutional field gradually become similar, a process often driven by coercive, mimetic, and normative pressures [2]. Scott, in turn, summarizes expectations within the institutional environment as three pillars: regulatory, normative, and cultural-cognitive [3]. Collectively, these studies demonstrate that legitimacy is not merely a matter of organizational efficiency; it also concerns whether an organization conforms to institutional rules, whether it can be understood by the outside world, and whether it can be regarded as a self-evident entity. Precisely for this reason, legitimacy can, to a certain extent, reduce an organization's uncertainty and vulnerability when facing the external environment [5].

This perspective was later further developed in entrepreneurship research. For startups, legitimacy is not an abstract theoretical label, but rather a prerequisite for being taken seriously by the outside world. Only when investors, customers, and partners are willing to view a startup as a genuine, credible, and trustworthy organization can it more easily secure resources and development opportunities [9–12]. Building on this, Suchman further emphasized organizational agency [4]. Legitimacy is not merely a result bestowed upon an organization by the external environment; organizations also actively seek, maintain, and repair their legitimacy. His distinction between instrumental, moral, and cognitive legitimacy remains explanatory to this day. Different audiences focus on different issues: some are concerned with whether the organization serves their own interests, others with whether the organization is morally acceptable, and still others with whether the organization is sufficiently clear, familiar, and easy to understand.

Research on strategic legitimation, however, examines the issue of legitimacy within more uncertain contexts. Ashforth and Gibbs note that when organizations defend their legitimacy too overtly, it may actually provoke external skepticism [6]. Elsbach demonstrates how organizations respond to controversies through explanatory statements [7]. Oliver discusses the various reactions organizations exhibit when facing institutional pressures, ranging from compliance to manipulation [8]. These studies also explain why the concept of legitimacy frequently appears in research on corporate social responsibility

communication, sustainability reporting, crisis response, public relations, and non-market strategies.

2.2 The Social Judgment Perspective on Legitimacy

Another line of research views legitimacy as a social judgment. In other words, legitimacy is not stored within an organization like an asset, but is bestowed upon the organization by external evaluators based on certain criteria [16–18]. These criteria may be cognitive, moral, instrumental, or relational. This perspective is significant because organizations in the real world rarely face only a single audience. Businesses, startups, multinational corporations, and digital platforms often simultaneously engage with investors, customers, regulators, the media, communities, and other stakeholders, and the judgment criteria of these actors are not always consistent.

Bitektine's distinction between legitimacy, reputation, and status offers valuable insights into this issue [16]. Legitimacy primarily concerns whether an organization is acceptable within a given social system; reputation involves comparative external evaluations of an organization's quality or performance; and status emphasizes an organization's relative position within a relational structure. Trust concerns external expectations regarding an organization's reliability or good faith, whereas credibility concerns whether a particular claim is believable. Although these concepts are frequently intertwined in empirical research, they cannot be simply equated. Otherwise, legitimacy risks being generalized into a catch-all term for nearly all positive evaluations [19].

In contemporary management research, audience diversity has become an unavoidable context for studies on legitimacy. Corporate social responsibility (CSR) and ESG research involve investors, non-governmental organizations (NGOs), the media, rating agencies, and regulators; international business research focuses on host-country governments and local communities; entrepreneurship research emphasizes investors, customers, and market intermediaries; and digital platform research further incorporates users, complementary parties, and algorithmic evaluators into the discussion. From this perspective, legitimacy is no longer a single judgment but rather a set of evaluations formed by different stakeholders based on their distinct concerns.

2.3 Research Gaps

The concept of legitimacy has been widely applied in research fields such as corporate social responsibility, environmental disclosure, sustainability reporting, and ESG. In accounting and disclosure research, scholars often use legitimacy to explain why firms disclose social and environmental information under external pressure [20–22, 26–28]. However, a persistent question remains: do these disclosures truly reflect the firm's genuine commitment to responsibility, or are they primarily used to maintain external acceptance of the firm? Similar discussions exist in international business research. When multinational enterprises enter different institutional environments, they face evaluation standards that are not entirely consistent, making them more susceptible to legitimacy pressures and adaptation issues [23–25].

Research on entrepreneurship and market categories has shifted the discussion in another direction. Before gaining recognition, startups and emerging market categories must first make themselves "understandable" to the outside world. Storytelling, identity statements, product prototypes, alliances, certifications, and category labels can all serve as signals of cognitive and instrumental legitimacy [9–12]. Digital organizations further complicate this issue. They must not only earn the trust of investors and users but also address the concerns of regulators and the public regarding privacy, fairness, transparency, and accountability [38–40]. These developments underscore the need for a bibliometric review to examine the diffusion of legitimacy research across different fields and how various research themes interconnect and reconfigure.

3. Data Sources and Research Methods

3.1 Data Sources and Search Strategy

The literature data in this study were obtained from the Web of Science Core Collection. This database was selected primarily because it provides relatively comprehensive and standardized bibliographic information and citation data, thereby meeting the requirements of bibliometric research, such as citation analysis, co-citation analysis, and keyword analysis [29,30]. Since this study focuses on research on legitimacy in the fields of business and management, the search scope was limited to journals indexed in SSCI and classified under the Business, Business Finance, and Management categories.

Regarding the selection of search terms, this study did not use narrow expressions such as "organizational legitimacy,"

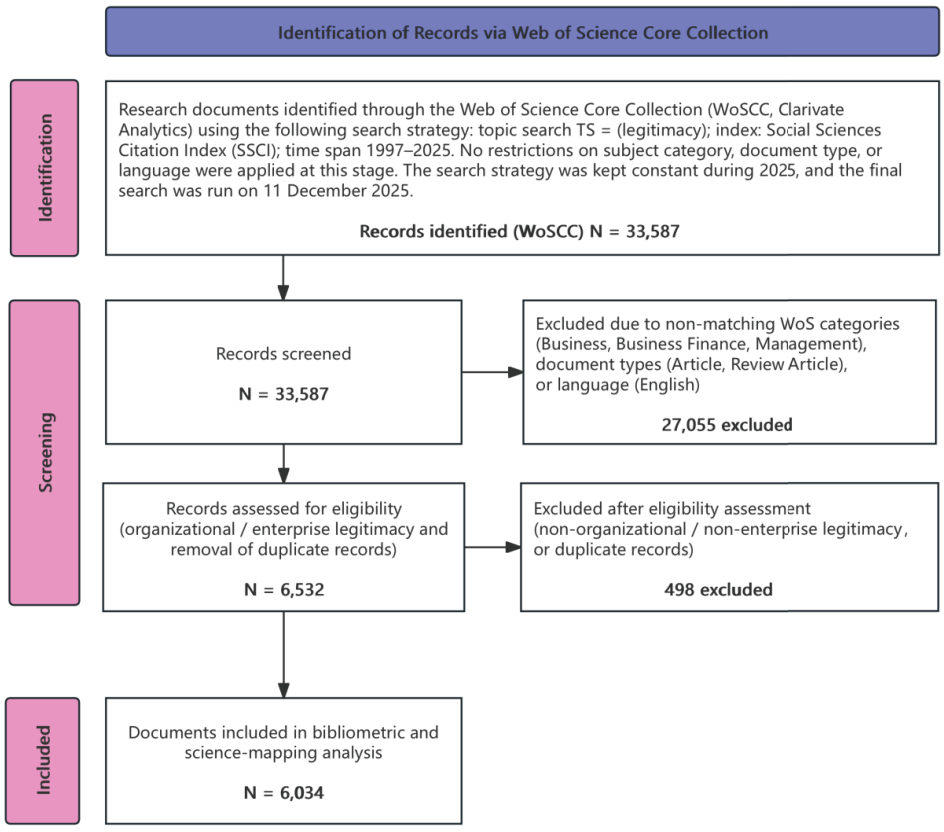
but instead employed a broader subject search formula: TS = (legitimacy). This study does not merely aim to examine research on organizational legitimacy in the narrow sense, but rather seeks to observe how the concept of legitimacy is used, disseminated, and reconfigured within business and management research. If narrower search terms had been used from the outset, the dataset might have been more concentrated, but it would also have been prone to omitting literature from other fields that utilize the concept of legitimacy. Of course, a broad search also carries a risk: not all articles included in the dataset treat legitimacy as a core theoretical concept. Some articles may use legitimacy merely as a supporting concept. Given this, this paper exercises caution in subsequent screening and interpretation of results, refraining from interpreting the dataset as a comprehensive collection of "organizational legitimacy research" but rather viewing it as a literature foundation for observing the diffusion of the legitimacy concept and the evolution of the topic.

The search period for this study spans from 1997 to 2025, with the final data extraction date set as December 11, 2025. The initial search yielded a total of 33,587 records. Subsequently, the study applied filters based on index type, disciplinary category, document type, and language, limiting the data to English-language SSCI journal articles and review papers. On this basis, duplicate records were further removed, and literature primarily unrelated to discussions of legitimacy within the business and management research context was excluded. Ultimately, the study yielded 6,034 valid literature records. It should be noted that some foundational literature published prior to 1997—such as studies by Meyer and Rowan, DiMaggio and Powell, Ashforth and Gibbs, Elsbach, and Suchman—were not included in the dataset as source documents but were nonetheless used as important references for theoretical interpretation in this study [1,2,4,6,7].

3.2 Literature Screening and Preprocessing

The literature screening process consisted of four main steps. First, records with the term "legitimacy" in the subject field were identified in Web of Science. Second, database filters were applied to retain literature within the business and management categories. Third, a manual review was conducted based on titles, abstracts, keywords, and source information to exclude literature primarily discussing political legitimacy, constitutional legitimacy, or legal legitimacy. Fourth, the full records and cited references of the remaining literature were exported for subsequent bibliometric analysis. Figure 1 illustrates the literature retrieval and screening process in this study.

Figure 1. Literature search and screening process



3.3 Analytical Procedures

This study combines descriptive statistics with scientific knowledge graph analysis. Descriptive statistics are primarily used to present the annual publication volume, document types, journal distribution, and basic characteristics of the datasets. Co-citation analysis is used to identify knowledge connections between documents; specifically, when a subsequent study cites two documents simultaneously, a co-citation relationship is formed between those two documents [33]. Keyword co-occurrence and clustering analysis are used to observe the concentration of research topics, while sudden emergence detection is employed to identify terms or references that have rapidly gained attention during specific periods [32].

This study uses CiteSpace to generate co-citation networks, keyword maps, clustering results, timeline charts, and burst detection results [32]. Descriptive charts were primarily compiled using Excel. To analyze the evolution of research themes, the study period is divided into four phases: 1997–2004, 2005–2012, 2013–2019, and 2020–2025. This division is not a strict historical periodization but rather an analytical framework established based on the development of the research. It is primarily used to observe the early diffusion of legitimacy research, its expansion into corporate social responsibility and entrepreneurship studies, its consolidation across different fields of management, and its recent reconfiguration in relation to themes such as ESG, sustainability, governance, and platforms.

The interpretation of the graph results is primarily based on the theoretical framework proposed in Section 2. Growth in the number of publications reflects the expansion of the field; co-citation patterns help elucidate the connections between research traditions such as institutional theory, strategic legitimation, social judgment, entrepreneurship, corporate social responsibility, and sustainability. Keyword clustering reflects shifts in research focus rather than direct changes in the theories themselves. Therefore, this paper does not treat bibliometric results as automatically generated conclusions but rather as a basis for further interpreting the evolution of legitimacy research [29, 30, 34].

Caution is also required when using citation metrics. Citation frequency and co-citation networks can help identify highly visible literature and knowledge proximity, but they do not directly represent theoretical quality [35, 36]. A frequently cited work may be so because it proposes a particular method, because its research context is relatively typical, or because it employs conceptual terminology that is easily adopted by other studies. Consequently, while this paper uses citation evidence as a starting point for analysis, it returns to the theory of legitimacy itself when interpreting the theoretical significance of these patterns.

4. Overview of Research on Legitimacy

4.1 Publication Patterns and Journal Distribution

The final dataset comprises 6,034 legitimacy-related publications in the fields of business and management from 1997 to 2025, including 5,803 research articles and 231 review articles. The dataset involves 11,273 authors, 2,703 institutions, 365 journals, and 47 Web of Science subject categories. The number of subject categories exceeds the number of search categories because many journals are classified under multiple categories simultaneously. This distribution demonstrates that the concept of legitimacy is no longer confined to a single discipline.

Table 1. Summary statistics of the dataset, 1997-2025

Category	Publication	Articles	Review Articles	Authors	Institutions	Journals	Subject categories
Amount	6034	5803	231	11273	2703	365	47

Figure 2 illustrates a process that began slowly but then accelerated significantly. Between 1997 and 2007, the annual number of publications was relatively low, rising from a small number of papers in the late 1990s to 63 in 2007. The first significant turning point occurred after 2008, when the annual output reached 116 papers and continued to rise, albeit with some fluctuations. The second shift occurred around 2018–2019, with output increasing from 327 to 411 papers. By 2024 and 2025, annual output is projected to reach 551 and 560 papers, respectively.

Figure 2. Annual publication output of legitimacy-related research in business and management, 1997-2025

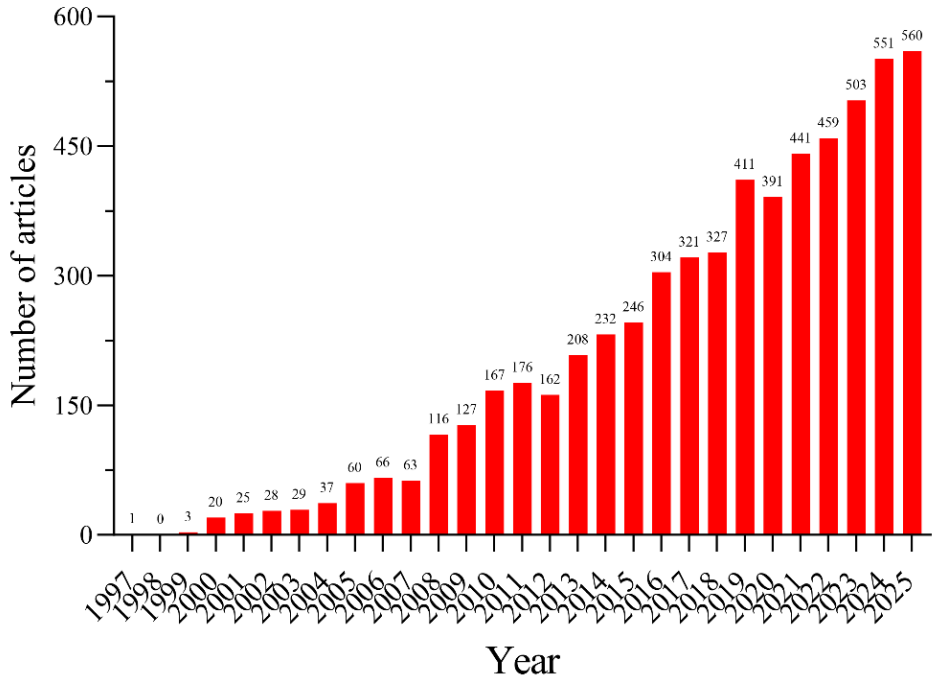
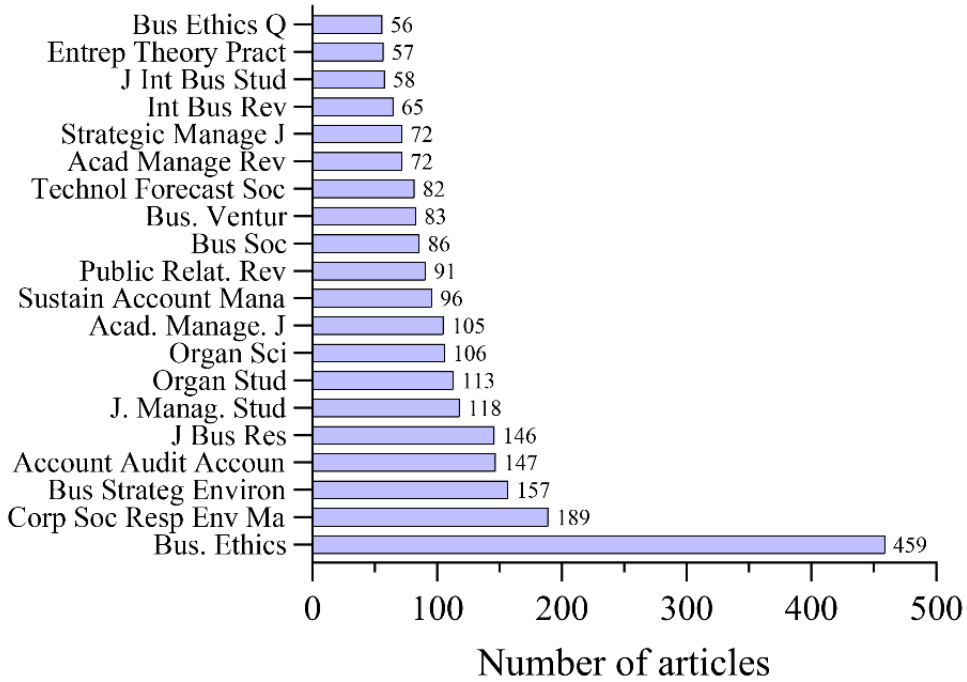


Figure 3 shows the journal distribution. Journal of Business Ethics is the most prominent outlet, with 459 related publications, followed by Corporate Social Responsibility and Environmental Management with 189 publications and Business Strategy and the Environment with 157 publications. This distribution is revealing.

Figure 3. Top journals publishing legitimacy-related research in business and management



4.2 Co-Citation Structure and Important Publications

This paper employs co-citation analysis to examine the knowledge structure of the selected literature. Figure 4 shows that the network comprises 1,793 nodes and 7,625 links, with a network density of 0.0049. The largest connected component contains 1,740 nodes, accounting for 97% of all nodes. This indicates that although the co-citation network of legitimacy research is relatively dispersed, most of the literature remains within the same knowledge-linked structure and does not constitute isolated research segments.

Looking at representative nodes in the figure, Scott's research demonstrates the enduring influence of institutional theory on legitimacy studies, suggesting that regulatory, normative, and cultural-cognitive institutional frameworks remain the foundation for understanding legitimacy. Freeman's stakeholder theory indicates that legitimacy research has increasingly focused on the relationship between organizations and external actors. Works by Bitektine, Suddaby, and Deephouse occupy prominent positions in the network, reflecting that social judgments, audience evaluations, and conceptual distinctions regarding organizational legitimacy have become key sources of knowledge in the field. Scherer's research links legitimacy to sustainable development and corporate responsibility, while studies by Taeuscher et al. indicate that the concept of legitimacy is extending into new contexts such as digital markets and startups. Overall, Figure 4 does not depict old theories being replaced by new themes, but rather the ongoing connections and reconfigurations among different research traditions.

Figure 4. Co-citation network of legitimacy-related research within the selected dataset

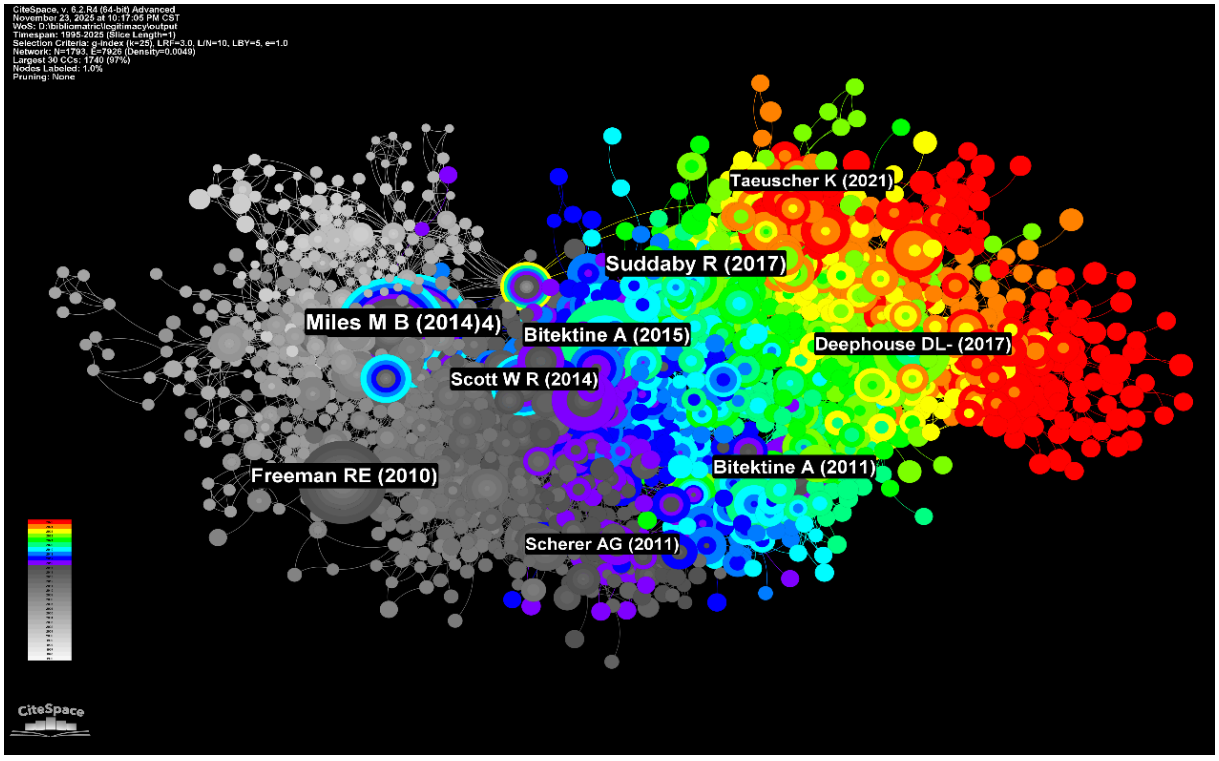


Table 2. Leading publications within the selected dataset (top 10 by local citation score)

No.	Publication	Journal	LCS	GCS
64	Beyond survival: Achieving new venture growth by building legitimacy	Academy of Management Review	534	1498
40	Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources	Strategic Management Journal	401	1383
460	How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship	Administrative Science Quarterly	304	921
1285	Toward a theory of social judgments of organizations	Academy of Management Review	249	808
2892	The macro and the micro of legitimacy	Academy of Management Review	236	609
3280	Legitimacy	Academy of Management Annals	206	510
205	Organizational legitimacy under conditions of complexity	Academy of Management Review	194	1161
802	Corporate legitimacy as deliberation	Journal of Business Ethics	171	764
1145	Managing legitimacy in complex and heterogeneous environments	Journal of Management Studies	157	575
2310	Legitimacy in organizational institutionalism	SAGE Handbook	142	760

Table 2 presents the publications with relatively high citation frequency in the selected dataset. In terms of research focus, these studies are distributed across different directions, including cultural entrepreneurship, market-category formation, social judgment, institutional complexity and corporate legitimacy.

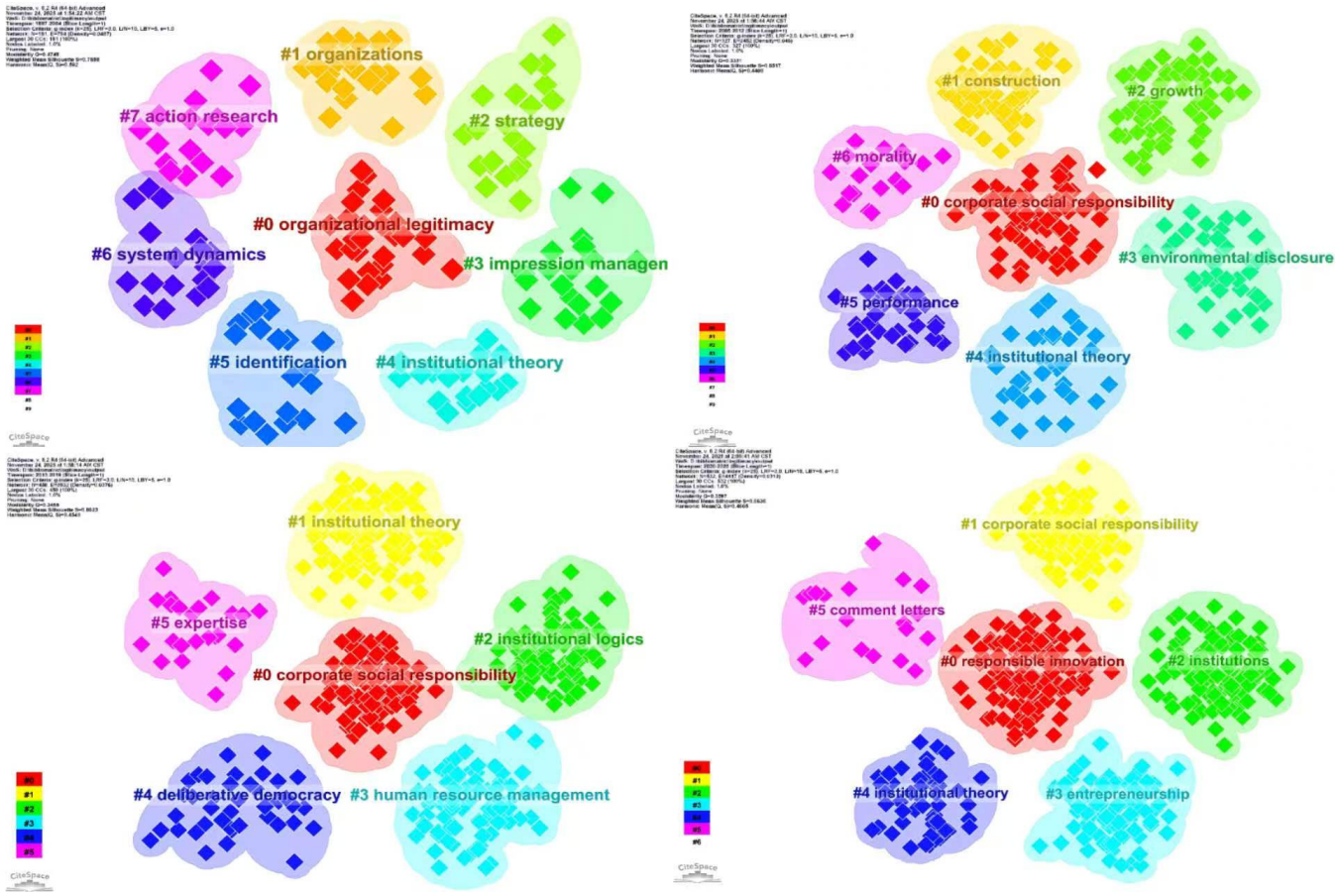
5. Research Hotspots and Thematic Evolution

5.1 Keyword Clustering and Research Hotspots

Figure 5 presents the results of keyword clustering across four phases from 1997 to 2025. Overall, the evolution of themes in legitimacy research shows a trend of expanding from institutional foundations to accountable governance and innovative contexts. The early stages primarily revolved around clusters such as "organizational legitimacy," "institutional theory," "strategy," and "impression management," indicating that research focused on how organizations obtain institutional recognition, manage their public image, and ensure their survival.

As research progressed, themes such as "corporate social responsibility," "environmental disclosure," "performance," "growth," and "morality" gradually came to the fore, suggesting that the concept of legitimacy began to be increasingly applied to explain the relationships among corporate social responsibility, environmental disclosure, and organizational performance. In the most recent phase, clusters such as "responsible innovation," "institutions," "entrepreneurship," and "comment letters" have become more prominent, indicating that legitimacy research is further expanding into new contexts such as innovation governance, entrepreneurial activities, and regulatory communication. Overall, Figure 5 demonstrates that institutional theory remains a crucial foundation for legitimacy research, but its application has clearly expanded into areas such as CSR, sustainability, entrepreneurship, and innovation governance.

Figure 5. Keyword clusters across four analytical periods, 1997-2025



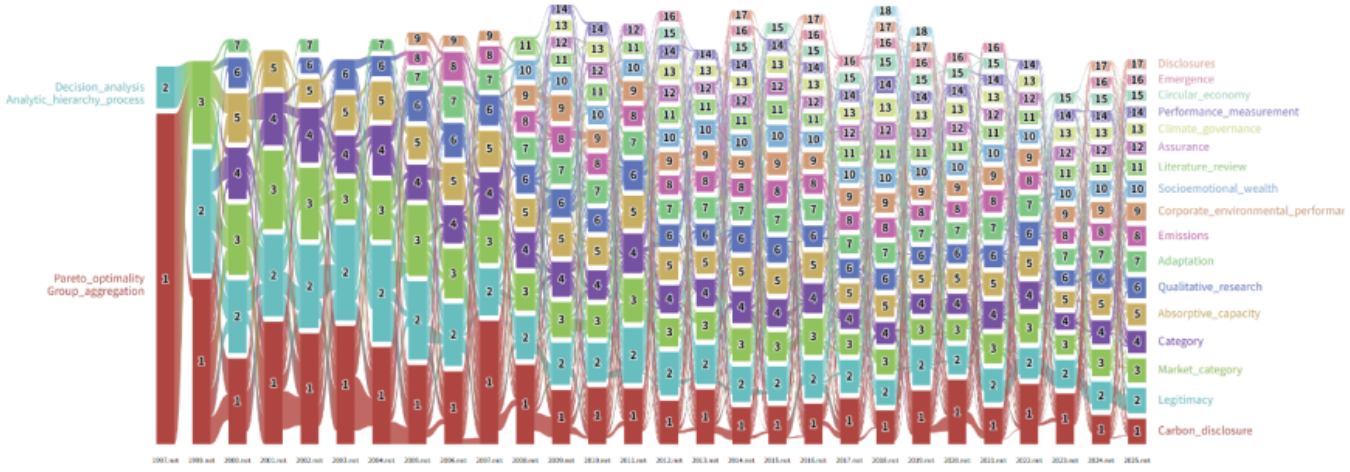
5.2 Thematic Reorganization and Alluvial Analysis

Figure 6 illustrates the annual shifts in research themes related to legitimacy between 1997 and 2025. Overall, early themes were primarily concentrated on methodological issues such as decision analysis, the analytic hierarchy process, Pareto

optimality, and group aggregation, indicating that early research in this area still contained a significant amount of content related to decision analysis and methodological applications. Subsequently, themes such as "legitimacy," "market category," "category," "qualitative research," and "adaptation" gradually stabilized, indicating that legitimacy research began to shift more toward the contexts of organizational studies, market categories, and qualitative research.

From the mid-to-late period onward, the flow of themes shifted noticeably toward environmental disclosure and sustainable governance. "Carbon disclosure" formed a relatively consistent backbone in the figure, while themes such as "emissions," "corporate environmental performance," "assurance," "climate governance," "circular economy," and "disclosures" also appeared frequently in the later stages. This suggests that the concept of legitimacy has not departed from the core issues of "recognition" and "acceptance" discussed in the early stages, but has instead been reapplied to explain how firms respond to external evaluations through carbon disclosure, environmental performance, sustainability reporting, and governance practices. Thus, Figure 6 does not depict a simple replacement of themes, but rather the continuous migration and reorganization of the concept of legitimacy across different management issues.

Figure 6. Alluvial diagram of thematic evolution in legitimacy-related research, 1997-2025

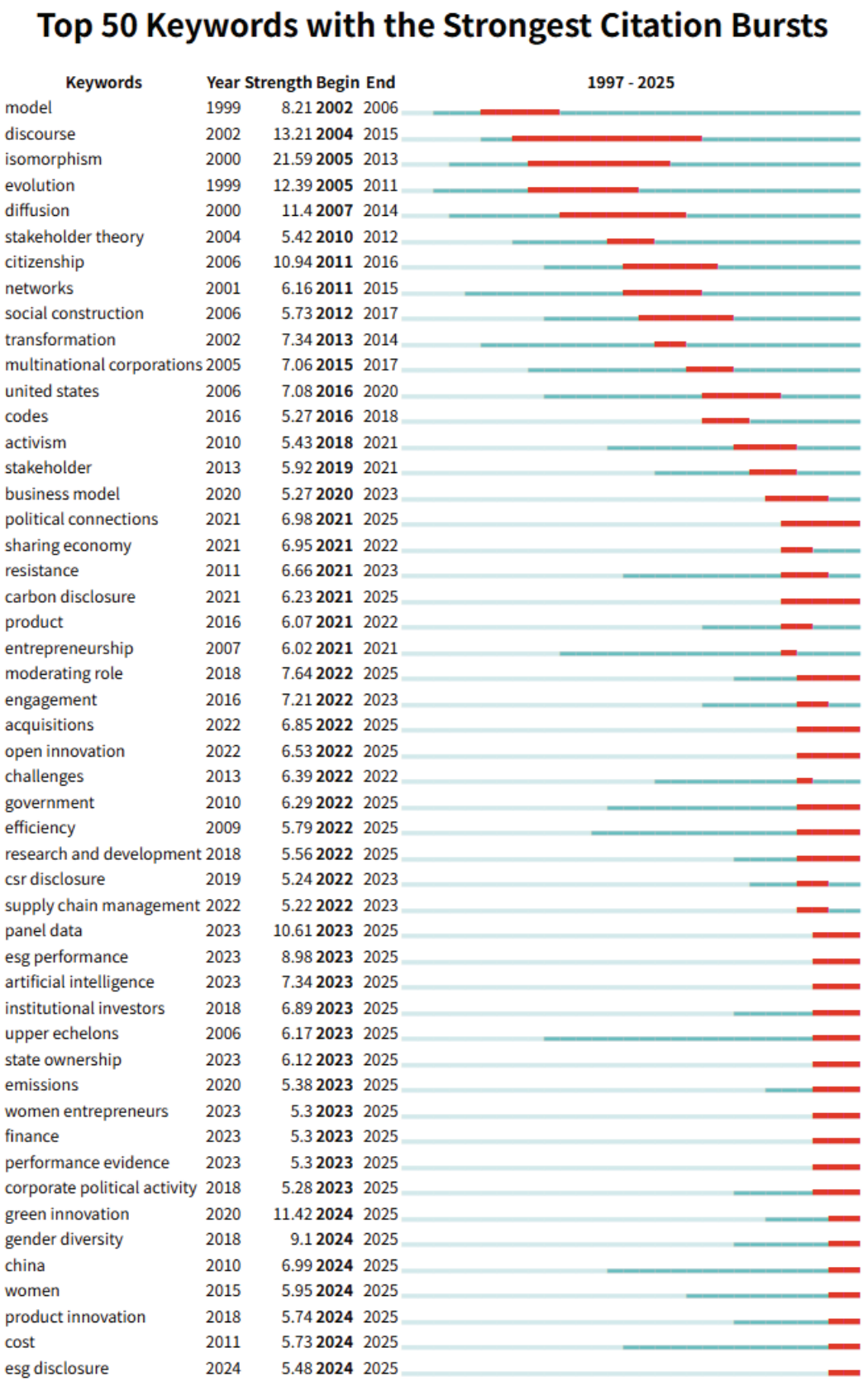


5.3 Research Trends and Frontier Directions

Figure 7 presents the keywords with the strongest citation bursts in legitimacy-related research from 1997 to 2025. In the early stage, burst terms such as model, discourse, isomorphism, evolution and diffusion were prominent. Among them, isomorphism shows the strongest burst strength, reaching 21.59, which indicates that institutional isomorphism was one of the most visible frontier topics in the earlier development of legitimacy research. The relatively long burst periods of discourse, evolution and diffusion also suggest that early studies were concerned not only with institutional conformity, but also with how legitimacy was constructed through discourse, change and diffusion.

In the middle and later stages, the burst keywords gradually shifted to stakeholder theory, citizenship, networks, social construction, multinational corporations and activism. This reflects a broader move from institutional conformity toward stakeholder relations, social construction, international business and social action. More recent burst terms show a stronger connection with ESG, green innovation and data-driven research. Panel data has a burst strength of 10.61, ESG performance 8.98, artificial intelligence 7.34, green innovation 11.42 and gender diversity 9.10, and these topics remain active until 2025. Overall, Figure 7 suggests that the frontier of legitimacy research has expanded from institutional isomorphism and discourse-based studies to carbon disclosure, ESG evaluation, artificial intelligence, institutional investors, green innovation and gender diversity, with increasing attention to quantitative measurement and empirical testing.

Figure 7. Top keyword bursts in legitimacy-related research



6. Discussion

6.1 Theoretical Continuity

Based on the findings discussed above, the expansion of legitimacy research is not a process in which old theories are replaced by new themes. Within the co-citation network, institutional theory, stakeholder theory, social judgment research,

and entrepreneurship research collectively form the knowledge foundation of this field. Keyword clustering and emerging terms also indicate that while legitimacy research has expanded into new topics such as CSR, environmental disclosure, ESG, green innovation, and digital platforms, its core issues remain unchanged. Organizations still need to be understood, recognized, and accepted by their external environment; however, this recognition no longer comes solely from traditional institutional settings but increasingly from diverse stakeholders such as investors, regulators, the media, consumers, communities, and platform users.

This also demonstrates that the value of the legitimacy concept lies in its strong interpretive flexibility. It can be used to explain how startups secure resources and market recognition, as well as to analyze how firms address external scrutiny through social responsibility disclosures, environmental reporting, and governance practices. However, the broader the concept is applied, the more easily its boundaries become blurred. If legitimacy is simply equated with reputation, trust, performance, credibility, or positive evaluations in a general sense, it loses its original analytical power [13–19]. Therefore, future research still needs to clarify more explicitly which type of legitimacy is being discussed, who makes the judgment, and what the basis for that judgment is.

6.2 Disclosure and Audience

The results of this paper indicate that CSR, environmental disclosure, carbon disclosure, ESG performance, and sustainability reporting have become highly active research areas within the field of legitimacy studies. In Figures 5 and 6, themes such as corporate social responsibility, environmental disclosure, carbon disclosure, climate governance, and sustainable governance have continued to gain prominence in the mid-to-late stages; in Figure 7, keywords such as carbon disclosure, ESG performance, green innovation, and ESG disclosure remain active through 2025. This indicates that legitimacy research is clearly moving toward issues of responsible governance and sustainable development.

These issues are significant because they vividly illustrate the differences in audiences within legitimacy assessments. When companies disclose social and environmental information, they may be presenting actual performance, or they may be responding to criticism, maintaining their image, or seeking social acceptance [20–22, 41–44]. The same disclosure may be evaluated differently by investors, regulators, the media, consumers, and activists. Investors may prioritize ESG ratings, activists may question whether companies are engaging in greenwashing, while regulators may focus on whether the disclosed information is sufficient and verifiable [48–55]. Therefore, legitimacy is not necessarily gained or lost as a unified whole; it may be differentiated across different audiences. Future research that merely asks whether a company "possesses legitimacy" would be overly general; more meaningful questions include who is making the judgment, on what basis, and how these judgments influence organizational behavior.

6.3 Directions for Future Research

Future research can proceed in three directions. First, legitimacy must be examined within a temporal process. It is not something that is acquired once and remains static; rather, it is constructed, eroded, repaired, or reinterpreted through events, controversies, and ongoing interactions. Changes in legitimacy are often particularly complex when judgments by investors, regulators, the media, communities, and users diverge. Longitudinal case studies, process tracking, and event studies can better capture these changes [17, 18, 37].

Second, the measurement of legitimacy requires greater caution. Media tone, certifications, awards, the volume of disclosures, ESG ratings, surveys, and investment outcomes can all serve as indicators for observing legitimacy, but they do not all point to the same thing. Certifications may be closer to regulatory or normative legitimacy, media evaluations may reflect public recognition, while investment outcomes may more closely reflect instrumental legitimacy. Theoretically discussing one form of legitimacy while empirically measuring another phenomenon weakens the explanatory power of research conclusions. Third, hybrid organizations and digital contexts warrant further attention. Social enterprises must strike a balance between social and business logics, while issues such as digital platforms, artificial intelligence, algorithmic governance, data privacy, and digital finance continually raise new questions regarding transparency, accountability, and the creation of social value [38–40, 45–59].

6.4 Research Limitations

This study has certain limitations. The research data primarily comes from the Web of Science Core Collection and is

restricted to English-language SSCI journals in the fields of business and management; consequently, books, book chapters, non-English literature, conference papers, and relevant research from adjacent disciplines have not been fully incorporated. The use of the search query TS = (legitimacy) facilitates the observation of the diffusion of the concept of legitimacy but may also include some literature that treats legitimacy merely as a secondary concept.

Furthermore, while bibliometric methods can demonstrate the visibility, interconnections, and thematic shifts in research, they cannot, on their own, elucidate the theoretical implications underlying these relationships [29, 30, 34–36]. Consequently, the network results presented here are best understood as structural clues rather than definitive explanations. Future research could build upon this foundation by integrating systematic content analysis to further examine how different studies define, utilize, and measure legitimacy.

7. Conclusion

Based on 6,034 literature records from the Web of Science database, this study examines research related to legitimacy in the fields of business and management from 1997 to 2025. The findings reveal that legitimacy research has expanded from its foundations in institutional theory and strategic studies to encompass multiple contexts, including entrepreneurship, CSR, sustainability reporting, ESG, governance accountability, international business, and digital platforms. Analysis of publication trends, journal distribution, co-citation networks, keyword clustering, keyword emergence, and thematic shifts indicates that new research themes have not supplanted the traditional framework of early legitimacy research but are instead continuously connecting and reconfiguring these traditions.

The concept of legitimacy remains important because it links organizational actions, social expectations, audience judgments, and the institutional environment. However, its flexibility also carries risks: if used too broadly, it can easily become merely another term for “approval” or “positive evaluation.” Future research needs to clarify the types of legitimacy, relevant audiences, evaluation criteria, and the processes by which legitimacy is acquired, maintained, challenged, restored, or lost. Only by clarifying these boundaries can legitimacy continue to serve as an explanatory analytical tool in business and management research.

Funding

This research was supported by the Faculty Research Start-up Fund Project of Guangzhou Xinhua University under the project “A Study on the Coupling Mechanism between Resource Bricolage and Legitimacy Construction of Nascent Social Enterprises from the Perspective of Ecosystem Theory” (Grant No. 2025KYQNSK32).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340-363. DOI: 10.1086/226550.

[2] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147-160. DOI: 10.2307/2095101.

[3] Scott, W. R. (2014). Institutions and organizations: Ideas, interests, and identities (4th ed.). SAGE.

[4] Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of Management Review, 20(3), 571-610. DOI: 10.5465/amr.1995.9508080331.

[5] Singh, J. V., Tucker, D. J., & House, R. J. (1986). Organizational legitimacy and the liability of newness. Administrative Science Quarterly, 31(2), 171-193. DOI: 10.2307/2392787.

[6] Ashforth, B. E., & Gibbs, B. W. (1990). The double-edge of organizational legitimation. Organization Science, 1(2), 177-194. DOI: 10.1287/orsc.1.2.177.

[7] Elsbach, K. D. (1994). Managing organizational legitimacy in the California cattle industry: The construction and effectiveness of verbal accounts. Administrative Science Quarterly, 39(1), 57-88. DOI: 10.2307/2393494.

[8] Oliver, C. (1991). Strategic responses to institutional processes. Academy of Management Review, 16(1), 145-179. DOI: 10.5465/amr.1991.4279002.

[9] Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond survival: Achieving new venture growth by building legitimacy. Academy of Management Review, 27(3), 414-431. DOI: 10.5465/amr.2002.7389921.

[10] Lounsbury, M., & Glynn, M. A. (2001). Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. Strategic Management Journal, 22(6-7), 545-564. DOI: 10.1002/smj.188.

[11] Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990-2005. Administrative Science Quarterly, 55(3), 439-471. DOI: 10.2189/asqu.2010.55.3.439.

[12] Fisher, G., Kotha, S., & Lahiri, A. (2016). Changing with the times: An integrated view of identity, legitimacy, and new venture life cycles. Academy of Management Review, 41(3), 383-409. DOI: 10.5465/amr.2013.0496.

[13] Deephouse, D. L., & Suchman, M. (2008). Legitimacy in organizational institutionalism. In R. Greenwood, C. Oliver, K. Sahlin, & R. Suddaby (Eds.), The SAGE handbook of organizational institutionalism (pp. 49-77). SAGE.

[14] Suddaby, R., Bitektine, A., & Haack, P. (2017). Legitimacy. Academy of Management Annals, 11(1), 451-478. DOI: 10.5465/annals.2015.0101.

[15] Deephouse, D. L., Bundy, J., Tost, L. P., et al. (2017). Organizational legitimacy: Six key questions. In R. Greenwood, C. Oliver, T. B. Lawrence, & R. E. Meyer (Eds.), The SAGE handbook of organizational institutionalism (2nd ed., pp. 27-54). SAGE.

[16] Bitektine, A. (2011). Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. Academy of Management Review, 36(1), 151-179. DOI: 10.5465/amr.2009.0382.

[17] Bitektine, A., & Haack, P. (2015). The macro and the micro of legitimacy: Toward a multilevel theory of the legitimacy process. Academy of Management Review, 40(1), 49-75. DOI: 10.5465/amr.2013.0318.

[18] Tost, L. P. (2011). An integrative model of legitimacy judgments. Academy of Management Review, 36(4), 686-710. DOI: 10.5465/amr.2010.0227.

[19] Deephouse, D. L., & Carter, S. M. (2005). An examination of differences between organizational legitimacy and organizational reputation. Journal of Management Studies, 42(2), 329-360. DOI: 10.1111/j.1467-6486.2005.00499.x.

[20] Deegan, C. (2002). The legitimising effect of social and environmental disclosures: A theoretical foundation. Accounting, Auditing & Accountability Journal, 15(3), 282-311. DOI: 10.1108/09513570210435852.

[21] Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. Accounting, Organizations and Society, 32(7-8), 639-647. DOI: 10.1016/j.aos.2006.09.009.

[22] Cho, C. H., Laine, M., Roberts, R. W., et al. (2015). Organized hypocrisy, organizational facades, and sustainability reporting. Accounting, Organizations and Society, 40, 78-94. DOI: 10.1016/j.aos.2014.12.003.

[23] Kostova, T., & Zaheer, S. (1999). Organizational legitimacy under conditions of complexity: The case of the multinational enterprise. Academy of Management Review, 24(1), 64-81. DOI: 10.5465/amr.1999.1580441.

[24] Marano, V., Tashman, P., & Kostova, T. (2017). Escaping the iron cage: Liabilities of origin and CSR reporting of emerging market multinational enterprises. Journal of International Business Studies, 48(3), 386-408. DOI: 10.1057/jibs.2016.17.

[25] Tashman, P., Marano, V., & Kostova, T. (2019). Walking the walk or talking the talk? Corporate social responsibility decoupling in emerging market multinationals. Journal of International Business Studies, 50(2), 153-171. DOI: 10.1057/s41267-018-0171-7.

[26] Palazzo, G., & Scherer, A. G. (2006). Corporate legitimacy as deliberation: A communicative framework. Journal of Business Ethics, 66(1), 71-88. DOI: 10.1007/s10551-006-9044-2.

[27] Scherer, A. G., Palazzo, G., & Seidl, D. (2013). Managing legitimacy in complex and heterogeneous environments: Sustainable development in a globalized world. Journal of Management Studies, 50(2), 259-284. DOI: 10.1111/joms.12014.

[28] Bansal, P., & Clelland, I. (2004). Talking trash: Legitimacy, impression management, and unsystematic risk in the context of the natural environment. Academy of Management Journal, 47(1), 93-103. DOI: 10.5465/20159562.

[29] Zupic, I., & Cater, T. (2015). Bibliometric methods in management and organization. Organizational Research Methods, 18(3), 429-472. DOI: 10.1177/1094428114562629.

[30] Donthu, N., Kumar, S., Mukherjee, D., et al. (2021). How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research, 133, 285-296. DOI: 10.1016/j.jbusres.2021.04.070.

[31] Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), 959-975. DOI: 10.1016/j.joi.2017.08.007.

[32] Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for Information Science and Technology, 57(3), 359-377. DOI: 10.1002/asi.20317.

[33] Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. Journal of the American Society for Information Science, 24(4), 265-269. DOI: 10.1002/asi.4630240406.

[34] van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84(2), 523-538. DOI: 10.1007/s11192-009-0146-3.

[35] Garfield, E. (1979). Citation indexing: Its theory and application in science, technology, and humanities. Wiley.

[36] Bornmann, L., & Daniel, H. D. (2008). What do citation counts measure? A review of studies on citing behavior. Journal of Documentation, 64(1), 45-80. DOI: 10.1108/00220410810844150.

[37] Freeman, R. E. (2010). Strategic management: A stakeholder approach. Cambridge University Press. DOI: 10.1017/CBO9781139192675.

[38] Taeuscher, K. (2019). Reputation and new venture performance in online markets: The moderating role of market crowding. Journal of Business Venturing, 34(6), 105944. DOI: 10.1016/j.jbusvent.2019.06.005.

[39] Saebi, T., Foss, N. J., & Linder, S. (2019). Social entrepreneurship research: Past achievements and future promises. Journal of Management, 45(1), 70-95. DOI: 10.1177/0149206318793196.

[40] Gillespie, N., Lockey, S., & Curtis, C. (2021). Trust in artificial intelligence: A five country study. University of Queensland and KPMG Australia. DOI: 10.14264/e34bfa3.

[41] Hahn, R., & Kuhnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. Journal of Cleaner Production, 59, 5-21. DOI: 10.1016/j.jclepro.2013.07.005.

[42] Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. Critical Perspectives on Accounting, 33, 59-78. DOI: 10.1016/j.cpa.2014.10.003.

[43] Boiral, O. (2013). Sustainability reports as simulacra? A counter-account of A and A+ GRI reports. Accounting, Auditing & Accountability Journal, 26(7), 1036-1071. DOI: 10.1108/AAAJ-04-2012-00998.

[44] Etzion, D., & Ferraro, F. (2010). The role of analogy in the institutionalization of sustainability reporting. Organization Science, 21(5), 1092-1107. DOI: 10.1287/orsc.1090.0494.

[45] Greenwood, R., Raynard, M., Kodeih, F., et al. (2011). Institutional complexity and organizational responses. Academy of Management Annals, 5(1), 317-371. DOI: 10.1080/19416520.2011.590299.

[46] Pache, A. C., & Santos, F. (2010). When worlds collide: The internal dynamics of organizational responses to conflicting institutional demands. Academy of Management Review, 35(3), 455-476. DOI: 10.5465/amr.35.3.zok455.

[47] Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012). The institutional logics perspective: A new approach to culture, structure, and process. Oxford University Press. DOI: 10.1093/acprof:oso/9780199601936.001.0001.

[48] Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. California Management Review, 54(1), 64-87. DOI: 10.1525/cmr.2011.54.1.64.

[49] Lyon, T. P., & Montgomery, A. W. (2015). The means and end of greenwash. Organization & Environment, 28(2), 223-249. DOI: 10.1177/1086026615575332.

[50] Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of

greenwashing. Organization Science, 27(2), 483-504. DOI: 10.1287/orsc.2015.1039.

[51] Christensen, L. T., Morsing, M., & Thyssen, O. (2013). CSR as aspirational talk. Organization, 20(3), 372-393. DOI: 10.1177/1350508413478310.

[52] Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835-2857. DOI: 10.1287/mnsc.2014.1984.

[53] Ioannou, I., & Serafeim, G. (2015). The impact of corporate social responsibility on investment recommendations: Analysts' perceptions and shifting institutional logics. Strategic Management Journal, 36(7), 1053-1081. DOI: 10.1002/smj.2268.

[54] Kolbel, J. F., Busch, T., & Jancso, L. M. (2017). How media coverage of corporate social irresponsibility increases financial risk. Strategic Management Journal, 38(11), 2266-2284. DOI: 10.1002/smj.2647.

[55] Clementino, E., & Perkins, R. (2021). How do companies respond to environmental, social and governance (ESG) ratings? Evidence from Italy. Journal of Business Ethics, 171(2), 379-397. DOI: 10.1007/s10551-020-04441-4.

[56] Mair, J., Mayer, J., & Lutz, E. (2015). Navigating institutional plurality: Organizational governance in hybrid organizations. Organization Studies, 36(6), 713-739. DOI: 10.1177/0170840615580007.

[57] Battilana, J., & Lee, M. (2014). Advancing research on hybrid organizing: Insights from the study of social enterprises. Academy of Management Annals, 8(1), 397-441. DOI: 10.5465/19416520.2014.893615.

[58] Nambisan, S., Wright, M., & Feldman, M. (2019). The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. Research Policy, 48(8), 103773. DOI: 10.1016/j.respol.2019.03.018.

[59] George, G., Merrill, R. K., & Schillebeeckx, S. J. D. (2021). Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. Entrepreneurship Theory and Practice, 45(5), 999-1027. DOI: 10.1177/1042258719899425.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on the Multidimensional Network Spatial Structure and Correlation Effect of Urban Agglomerations：Analysis of China's Five Major Urban Agglomerations

Tingying Peng[1]*, Xiaoyue Huang[2]

1. Nanchang Normal University, Nanchang, Jiangxi, 330032, China
2. Macau University of Science and Technology, Macau 999078, China

Corresponding author: Tingying Peng, pty198801@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Because the flow of regional factors tends to be more diverse and complicated, strengthening research on the counterpoint network of urban agglomerations is highly important for promoting the high-quality development of urban agglomerations. Using the theory of regional growth, a theoretical framework of multidimensional network spatial connections among urban agglomerations is constructed. This paper analyzes the multidimensional network spatial structure and correlation effect of urban agglomerations in China. The results demonstrate that (1) the multidimensional network of urban agglomerations presents significant spatial differences in terms of a "point–line network". (2) The networks present a multithreaded, multicity, and cross-regional distribution pattern and exhibit "local area network" and "small world" characteristics. (3) The connections among commuting, capital, technology, and information have a significant positive effect on economic connections in urban agglomerations. Moreover, the Beijing–Tianjin–Hebei region and the middle reaches of the Yangtze River rely more on technology, the Yangtze River Delta relies more on commuting, and the Pearl River Delta and Chengdu–Chongqing rely more on information.

Keywords: Urban Agglomeration; Multidimensional Connection Network; Spatial Structure; Economic Connection; QAP Regression

Published: Jun 24, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1454

1. Introduction

With the rapid development of information technology and transportation, the connections between cities are becoming more complicated and diversified. The flow of various elements is intertwined, and networks are becoming increasingly diverse; in particular, the dynamic network of transportation, population, enterprise, information and technology has become the focus of attention (Pflieger & Rozenblat, 2010; Tirthankar et al., 2023). Research on urban flow spatial networks has gradually overcome the limitations of distance and physical flow, and more attention has been given to the network structure and virtual flow between regions. In this context, the big data of spatial locations and social media services provide rich, detailed and real-time information for urban network research, reducing the cost of data acquisition for and facilitating the systematic and comprehensive development of research on urban flow spatial networks (Zhen et al., 2019). Because diversified data have stimulated innovations in methods, the research content and scale of urban flow spatial networks have deepened (Liu et al.,

2012). Therefore, regional network connection research is complex and diverse. In the future, urban agglomerations will be an important force for promoting the coordinated development of regions. The urban pattern of coordinated development between large, medium, and small cities and small towns, with urban agglomerations as the main body, is developing in China. The study of the multidimensional network structure and correlation effect of urban agglomerations could enrich the existing research on urban agglomeration networks and expand the research on the effect of the spatial network of urban flows, contributing to the study of the interaction and impact of urban agglomerations as complex economic, cultural, administrative, and temporal distance systems. The study of multidimensional network structures and associated effects could effectively compensate for the problem of insufficient identification of the spatial structure of urban agglomerations through a single network study, providing city planners with insights into how to optimize the internal connectivity of urban agglomerations and improve their overall efficiency and competitiveness. By analyzing the transportation network, information flow, and economic linkages of urban agglomerations, planners could better design urban infrastructure and efficiently allocate and utilize resources.

Western scholars first deconstructed the logic of urban networks from the perspective of flow space, arguing that flow space is the material organization of social practices of shared time that operates through mobility (Castells, 1996). They proposed that complex networks of flows linked by cities as nodes in the context of globalization and informatization should be the point of departure for understanding modern urbanization (Pflieger et al., 2010). Since the 1990s, along with the reunderstanding of space in philosophy and the study of spatial economic activities in geography, there has been a gradual realization that space is constructed by a series of economic and social relations, leading to the development of the "Relationship Turn" in the field of economic geography (He et al., 2008). Bathelt & Glückler (2003) believe that the research elements of relational economic geography include organization, evolution, innovation, and the interaction of behavioral themes, emphasizing the heterogeneity of regional network relations. In terms of network types, scholars have focused mainly on production networks (Derudder et al., 2010; Xu et al., 2024; Yang, 2024), innovation networks (Huggins & Prokop, 2017; Gui et al., 2019), the internet (Jacobs et al., 2011; Malecki, 2002a), and transportation networks (Malecki, 2002b; Yang et al., 2019; Tenkanen et al., 2023). In terms of spatial scale, geographic network research is divided into global and regional networks. Global network research includes production (Brown et al., 2010), innovation (Bathelt et al., 2004; Ma et al., 2023), and trade networks (Li, 2014), while regional network research comprises innovation, transportation, and capital connection networks. The theoretical methods mainly include gravity (Jung et al., 2008) and membership network models (Peter et al., 2010). These studies provide the theoretical and methodological basis for subsequent urban multidimensional network analysis.

The networks between regions are complex and diverse. In recent years, with the advancement of big data technology and network analysis methods, some scholars have begun to pay attention to the relationships among multidimensional networks across regions. Urban networks are diverse and include "multidirectional flows" of economic, social, cultural, and environmental activities (Davoudi, 2008). Different places are functionally connected not only through labor market relationships but also through trade, capital flows, consumer travel, school commuting, social visits, and leisure travel. This leads to a complex network of interregional relationships, where cities and regions rely on other cities and regions for education, labor, capital, intermediate products, and product markets (Burger, 2013). Zhang et al. (2019) analyzed the spatial and topological structure of the Yangtze River Delta’s urban agglomeration network using transportation, enterprise, population, and technology data, positioning this research framework as an empirical exposition of the formation of a diverse urban network. Urban networks are typically diverse with different forms of spatial interaction between cities, including spatial interactions among people, objects, and information (Hu et al., 2020). Moreover, Zhang et al. (2019) analyzed the multicity network in the Yangtze River Delta region using transportation, enterprise, population, and technology data. Therefore, a single-element network can no longer accurately describe the network relationships between regions, and research on multidimensional regional networks has gradually become a focus.

Huggins & Thompson (2014) proposed the concept of "network capital" and noted that the network relationship established by conscious investment is an important part of regional assets. The notion of network capital has been increasingly discussed and developed in the literature in recent years (Huggins & Johnston, 2010; Kramera et al., 2011), providing a useful and

necessarily specific means of conceptualizing and incorporating the asset value of investments in interorganizational networks into growth models. The network metaphor has become highly prevalent in research on economic geography and regional economic development, mostly related to innovation performance and economic development (Huber, 2012), which highlights the relevance of networks to economic growth and performance. Although the network concept is often applied broadly, few studies have sought to apply it in a more formal and constrained way, which would allow for a more precise rendering of the potential role of networks in regional growth processes. Therefore, drawing on the concept of network capital and incorporating flow space theory, this paper constructs a formal model to capture the network value of these processes.

Given the above theoretical and practical background, this paper makes the following three main contributions: First, on the basis of the theory of "flow space", a theoretical analytical framework is constructed for the multidimensional flow space network of urban agglomerations, which integrates the relevant theories of regional economic growth and spatial structure and reveals the mechanism of multidimensional elemental flow networks on economic linkage networks. Second, the multidimensional flow spatial networks of the five major urban agglomerations in China are established using data on traffic, information, science and technology, and capital flows, and their differential characteristics are comparatively analyzed via social network and GIS spatial analysis methods. In addition, the QAP analysis method is used to empirically test the promotion effect of the multidimensional flow spatial network of the five major urban agglomerations on the static network and to compare the differences in the network role relationships of different urban agglomerations.

The first part is described above. The second section introduces the theoretical mechanism. The third section describes the research methods and data sources. The fourth and fifth parts present the multidimensional network spatial structure and correlation effect of urban agglomerations, respectively. The sixth section presents the conclusions and policy implications.

2. Scoring Efficiency

2.1. Theoretical Mechanism of Network Growth

Economic growth rates are increasingly considered to depend on endogenous factors, with most treatments commonly assuming that economic growth is partly a function of either human capital (Lucas, 1988) or R&D (Romer, 1986). To better analyze the relationships and mechanisms of the multidimensional network structure of urban agglomerations, this paper conducts an in-depth analysis of the theoretical mechanism of urban agglomeration network growth using the basic theory of regional growth and the Cobb–Douglas production function model.

According to the Cobb–Douglas production function theory, there is a certain relationship between the number of production factors used in production and the maximum output that can be produced within a certain period of time, and the technology level remains unchanged. The specific model is as follows:

$$Yi=FKi,Li=KiaLia,0 < a < 1 \quad (1)$$

Here, suppose a country or city agglomeration has i cities, where Ki and Li represent the capital and labor invested by city i, respectively, and where Yi represents the production output of city i. Therefore, urban economic growth mainly depends on improvements in labor productivity, which involves improvements in factor allocation efficiency and increases in knowledge stock (Romer, 1986; Aghion & Howitt, 1998). With the development of science and technology, innovation has become an important force for promoting economic growth (Xu & Zhao, 2020). Some scholars have confirmed that the information development level represented by the digital economy also significantly promotes regional economic growth (Beckman et al., 2012).

Although these theories are still developing, there is little scholarly understanding of the mechanisms by which factors are transferred between organizations (Storper & Venables, 2004), which highlights the need for a research agenda that identifies the role of factor mobility in the spillover channel that promotes economic growth (Audretsch & Feldman, 1996). The production function theory has been introduced into the field of urban agglomeration network connections, and the network connections between cities are believed to have multiple attributes. The economic connection network is affected by financial, commuting, information, and technological connections. According to the neoclassical regional economic growth theory characterized by diminishing marginal returns, the self-accumulation and external inflow of capital, labor and technology are emphasized in the promotion of economic growth, with the interregional flow of production factors regarded as a critical factor driving economic growth. When there is no exogenous technological progress, the economic growth of various

countries or regions will converge (Taylor & Williamson, 2006; Justin et al., 2004). Population mobility has a significant effect on regional development (Barro & Sala-i-Martin, 1992). Capital flows could promote the reduction or expansion of regional economic differences (Jiao et al., 2021). Information flows reflect the economic and social connections between cities and a city's comprehensive strength to a certain extent (Zhen et al., 2013). Yao et al. (2001) noted that the association of high-tech enterprises could promote the formation of urban networks, enhance regional innovation competitiveness, and promote the coordinated development of cities. The model can be characterized as follows:

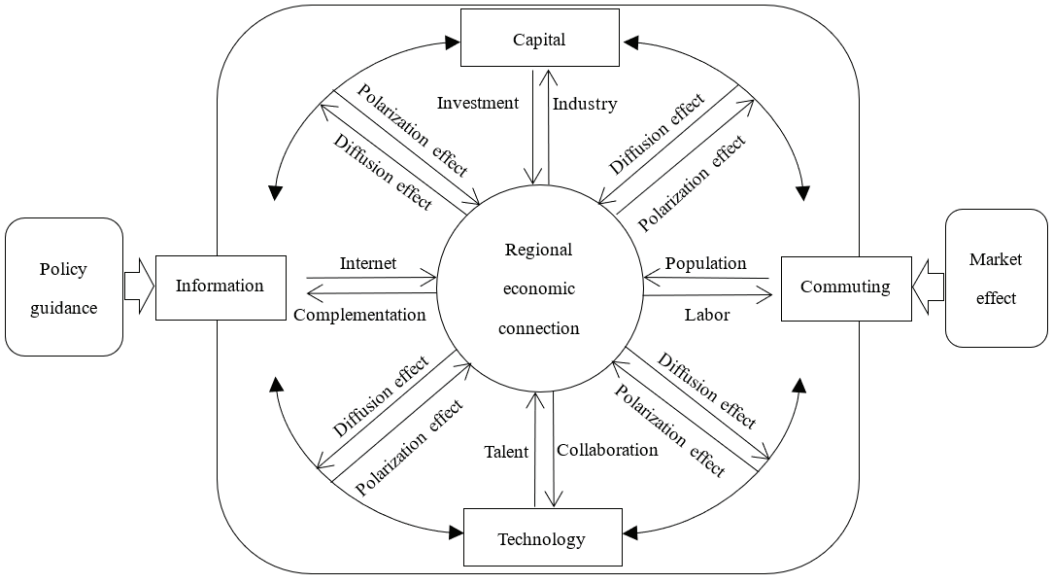
$$Y_{ij}=（FK_{ij},L_{ij},T_{ij},I_{ij}） \quad (2)$$

Assuming that a city cluster has several cities, Y_{ij} is the economic connection between cities i and j, and K_{ij}, L_{ij}, T_{ij}, and I_{ij} represent the capital, commuting , technology, and information connections between cities i and j, respectively. Capital connections refer to the collection of capital flows between cities, which could promote regional development and reduce regional differences by establishing regional industrial linkages and strengthening the intensity of regional linkages (Qian & Sun, 2021). Commuting flow refers to the intensity of population movements between cities, which could have an impact on the economic connections between cities by accelerating or slowing the population flow between cities. Technological connection refers to the strength of the relationship between the production and transformation of knowledge between cities, which could strengthen the economic connections between regions by deeply embedding knowledge in global or regional production networks. Information connection refers to the strength of network information connections between cities, which helps reduce the impact of urban agglomeration spatial distance and promote more coordinated and high-quality development of the economic connection network of urban agglomerations. Moreover, spatial agglomeration in different industries is a result of different driving mechanisms (Bing, 2020), and urban agglomeration network growth also has its own driving mechanism.

2.2 Structural Mechanism of Network Connections

According to the theoretical mechanism of network growth constructed in this research, the economic connection network is at the top of the model. The four networks of capital, commuting, technology, and information connections are located at the next level of the model. Among them, the diffusion and polarization effects of large cities to small and medium cities run through the connections between the various networks. Because the market plays a dominant role in the flow of factors, the market is in the core dominant position. For example, the density of urban networks in the developed eastern regions of China is greater than that in the western regions. In addition, policy guidance has a moderating effect on network connections. Regional development strategies promote cooperation within urban agglomerations, the flow of elements, and the development of urban agglomeration networks. For example, urban agglomerations include the Beijing–Tianjin–Hebei region, the Yangtze River Delta, the Guangdong–Hong Kong–Macao Greater Bay Area, the middle reaches of the Yangtze River, and Chengdu–Chongqing. The specific structural relationships are shown in Figs. 1.

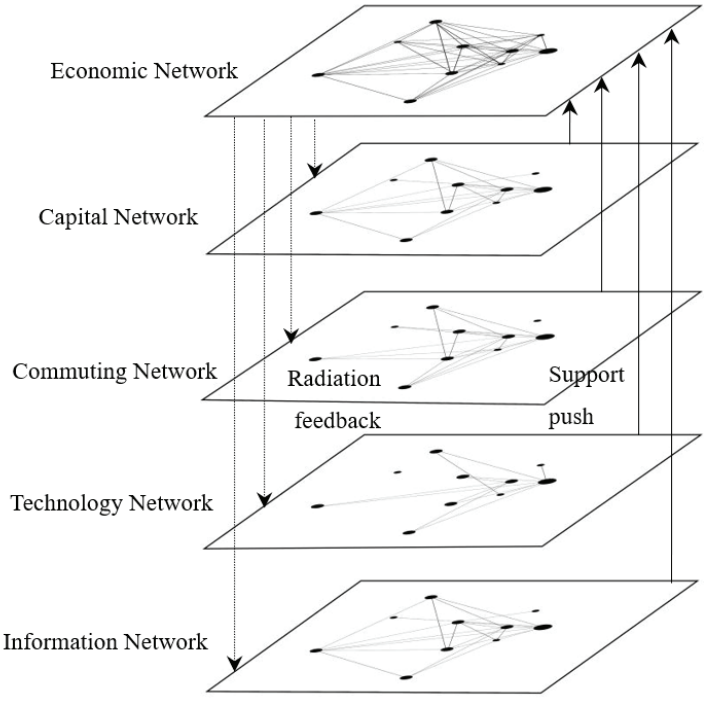
Figure 1. The multidimensional network structure of regional economic connections



2.3 The Mechanism of Network Interaction

The multilayer network of the economic connections of urban agglomerations is interdependent and interactive, which is manifested in the spatial relationship of agglomeration and diffusion. The connection networks of capital, commuting, technology, and information support and promote the economic connection network. In turn, the economic connection network radiates feedback on the other connection networks (Fig. 2).

Figure 2. The superposition mechanism of multidimensional connection networks of urban agglomerations



On the one hand, given the role of support and promotion, the most direct role of capital connections between cities is to increase the degree of industrial relevance between regions. Shared capital connections between cities are important in promoting regional development and reducing regional differences (such as parent companies setting up branches in different cities or mergers and reorganizations) (Qian & Sun, 2021). Therefore, the agglomeration and diffusion of capital in urban agglomerations brought about by enterprise investment have important impacts on promoting economic connections between cities. Commuting connections between cities represent the inter-city flow of human capital, which is not only related to the development gradient between regions but also to urban industrial connections and business exchanges. Therefore, commuting connections between cities accelerate the flow efficiency of people, logistics, and capital flows, thereby promoting the improvement of urban economic connections. Specifically, the number of trains is characterized by urban commuting connections, which can further deepen our understanding of the evolution of the spatial connection pattern and regional integration of urban agglomerations supported by high-speed rail networks (Zhang et al., 2020). Technological connections are the endogenous force driving the growth of the economic connection network between cities. With rapid technological innovation, increasingly frequent knowledge flows, and increasingly close innovation cooperation, the network theory of global production emphasizes external shocks such as cross-border production organizations and technological innovation networks. Therefore, knowledge spillovers between regions should be fully exerted to form a strong spillover relationship and innovation field to drive the high-quality integrated development of regional collaborative innovation and economic linkage networks. The information connection represents the transience and effectiveness of the internet, which could make the geographic space become more homogeneous. Therefore, information connections weaken the laws of spatial distance attenuation, promote connections between center and periphery cities, and accelerate the coordinated development of the economic connection network of urban agglomerations.

On the other hand, the radiation feedback effect of the urban economic connection network on other networks occurs mainly

through market regulation and government policy formulation. For example, strengthening regional economic connections in the Yangtze River Delta would encourage the government to strengthen the construction of infrastructure, such as internal high-speed rail and information networks, which would improve commuting and information connection networks within urban agglomerations. In addition, improvements in economic connections promote the flow of capital, technology and other factors, which manifest mainly in the transfer of industries and technology from central cities to small and medium-sized cities. As a result, the multidimensional connection network between cities is superimposed on the radiation feedback. For example, because the economic connections within the Pearl River Delta urban agglomeration are relatively high, the policies formulated by Guangdong to help promote the construction of the Guangzhou–Shenzhen Science and Technology Corridor would effectively improve the regional science and technology network and other network connections.

3. Research Methods and Data Sources

3.1 Overview of the Study Area

This paper studies the five major urban agglomerations in China: Beijing–Tianjin–Hebei, the Yangtze River Delta, the Pearl River Delta, Chengdu–Chongqing, and the middle reaches of the Yangtze River. This quadrilateral also includes two isosceles triangles: one is Beijing–Shanghai–Chengdu, where the distance from Chengdu to Beijing and from Chengdu to Shanghai is approximately 1500 kilometers; and the other is Shanghai–Guangzhou–Chengdu, where the distance from Guangzhou to Chengdu and from Guangzhou to Shanghai is approximately 1200 kilometers. The five urban agglomerations are exactly "diamond shapes". The five major urban agglomerations are the four pivot points and the intersection of the diagonal lines of the "diamond shape" structure. According to the Communiqué of the Seventh Population Census of All Regions of China in 2020, these five urban agglomerations constitute approximately 37.14% of the country's resident population. The five urban agglomerations constituted nearly 52.52% of the country's GDP in 2019 and are the core carriers of China's urbanization and economic growth, as well as the urban agglomerations with the most active factor flows and the highest degree of networking in China. The five urban agglomerations have polycentric, monocentric, and bicentric spatial structures, representing the spatial organizational structure characteristics of most of the cities in China. Research on the multidimensional network structure within the five urban agglomerations could provide references and lessons for identifying the network structure of other urban agglomerations as well as for planning and managing urban agglomerations.

3.2 Research Methods

3.2.1 Gravity model methods

The traditional gravity model measures the strength of economic connections between two cities on the basis of their GDP, population, and distance, with relatively few indicators measuring urban quality. This paper improves the traditional gravity model by comprehensively using GDP, population, fixed asset investment, and total retail consumer goods to characterize city size and choosing the shortest road distance between cities to represent the distance between cities. The specific model is as follows:

$$T_{ij}=\frac{\sqrt[4]{G_i\times P_i\times I_i\times S_i}\times\sqrt[4]{G_j\times P_j\times I_j\times S_j}}{D_{ij}^{\gamma}} \tag{3}$$

where T_{ij} is the degree of connection between city i and city j; G, P , I and S represent the GDP, population, fixed asset investment and total retail consumer goods of the city, respectively; D_{ij} represents the shortest road travel distance between city i and city j; and the coefficient of friction is generally taken as 2 (Lu et al., 2021).

3.2.2 Social network analysis

This paper introduces social network analysis to reveal the network structure and network correlation characteristics of the urban economic flow space. Social network analysis is an interdisciplinary analysis method for "relational data" (Ducruet & Beauguitte, 2014) that takes "relationships" as the basic unit and describes them using graph theory tools and algebraic model technology. The social network analysis tool used systematically describes the network structure characteristics of economic, capital, commuting, technological, and information connections in China's five major urban agglomerations. The quadratic assignment procedure (QAP) of social network analysis is used to analyze the supporting and promoting effects of capital, commuting, technology, and information connections on economic connections.

(1) Network density. Network density characterizes the closeness of overall urban connections. The formula for calculating the overall tightness of urban agglomeration networks via social network analysis is as follows (Ducruet & Beauguitte, 2014):

$$D^K=2L/n(n-1),K \tag{4}$$

The value of K is between 0 and 1. The closer the value of the network density DK is to 1, the more cities there are in the network, and the closer the relationships are. The maximum density in reality is 0.5 (Liu et al., 2017).

(2) Average path length. The average path length (L) refers to the number of edges in the shortest path connecting two points in all connected paths (i,j), which characterizes the degree of node separation in the network.

$$L=\frac{2}{n(n-1)}\sum_{i\neq j} d_{ij} \tag{5}$$

L refers to the average value between two nodes in the network, and d_{ij} represents the shortest distance from node i to node j (Camagni, & Salone, 1993).

(3) Average clustering coefficient. The ratio of the number of edges connected between all adjacent nodes to the maximum possible number of edges is used to measure the local clustering of network nodes (Cao et al., 2023).

$$C=\frac{1}{n}\sum C_i=\frac{1}{n}\sum \frac{2E_i}{k_i(k_i-1)} \tag{6}$$

where C is the average of the cluster coefficients of all nodes and E_i and k_i are the actual and possible maximum numbers of edges of node i, respectively (Cao et al., 2023).

(4) Network efficiency

$$E=\frac{1}{n(n-1)}\sum \frac{1}{d_{ij}} \tag{7}$$

where E is the reciprocal average of the shortest distance between two nodes in the network and where d_{ij} is the shortest distance between cities i and j in the network.

(5) QAP regression analysis. To examine the support or impact of commuting, capital, technological, and information networks on economic connections, it is necessary to conduct a correlation or regression analysis between networks. The quadratic assignment procedure (QAP) method converts the matrix of dependent and independent variables into a long vector matrix and compares the lattice values of the corresponding network nodes. In addition, using the similarity coefficient between matrices as the regression result can avoid the multicollinearity problem in traditional econometric regression analysis. The significance test of similarity coefficients adopts the random permutation method of row and column data. Using the multiple row and column permutations of matrix data, nonparametric tests are conducted on the coefficients and R^2, including matrix correlation, local array joint table, and multiple regression analyses. The specific expression is as follows:

$$m\cdot\ y=\beta+\sum_r \beta_r (m\cdot\ x_r)+\sum_s \beta_s (m\cdot\ x_s)+v \tag{8}$$

where $m \cdot y$ is the matrix of the connection network between different attributes in the sending and receiving relationships and is the comprehensive economic connection theory matrix. $m \cdot x_r$ represents the attribute relationship matrix formed by the attribute characteristics of the city in the r-dimension, which are commuting, capital, technological, and information connection network matrices. $m \cdot x_s$ represents the covariate matrix of network relationships between cities in the s dimension. β, β_r , and β_s are the parameters to be estimated, and v is the residual.

3.3. Matrix Construction and Data Sources

(1) Economic network.A gravity model is used to calculate the strength of economic connections between cities. The specific indicators include GDP, the permanent population, investment in fixed assets, and social retail consumer goods. The data were obtained from the "China City Statistical Yearbook (2019)".

（2）Capital network. With the nonlocal investment of enterprises representing interregional capital connections, the database of listed companies and holding subsidiaries of the CSMAR is used to match the cities where the parent company and the subsidiaries are located according to the direct holding ratio and registered capital of the listed companies and clearing data on holding relationships outside of urban agglomerations. Finally, the paper constructs the capital connection matrixes of the five major urban agglomerations.

（3）Commuting network. High-speed rails are important for assessing cross-city commuting within urban agglomerations. To more accurately represent the intensity and scale of commuting, the frequency of intercity HSR trains is used to measure the intensity of intercity commuting. The 12306 platform can be used to query high-speed rail (trains with G, D and C

prefixes) data between two cities, eliminate duplicate trains, and obtain data matrix data for high-speed rail trains in five major urban agglomerations.

（4）Technology network.The current common academic approach to science and technology linkage networks is to jointly file patents or publish papers. The number of patent collaborations is used to measure the scientific and technological connections between cities. Using the smart bud patent query website, data on national patent cooperation between cities in 2019 were obtained, and the scientific and technological cooperation matrix data of the five major urban agglomerations were obtained via data cleaning and screening.

（5）Information network.The attention given by residents of one city to another on cell phones and PCs is calculated using a large amount of data from the Baidu web and news search. Network attention is used to characterize the information connections. The geographical comparison function of the Baidu Index website is used to construct the information connection matrix of the five major urban agglomerations between cities.

4. Multidimensional Spatial Network Structure of Urban Agglomerations

4.1 Spatial Network Differences

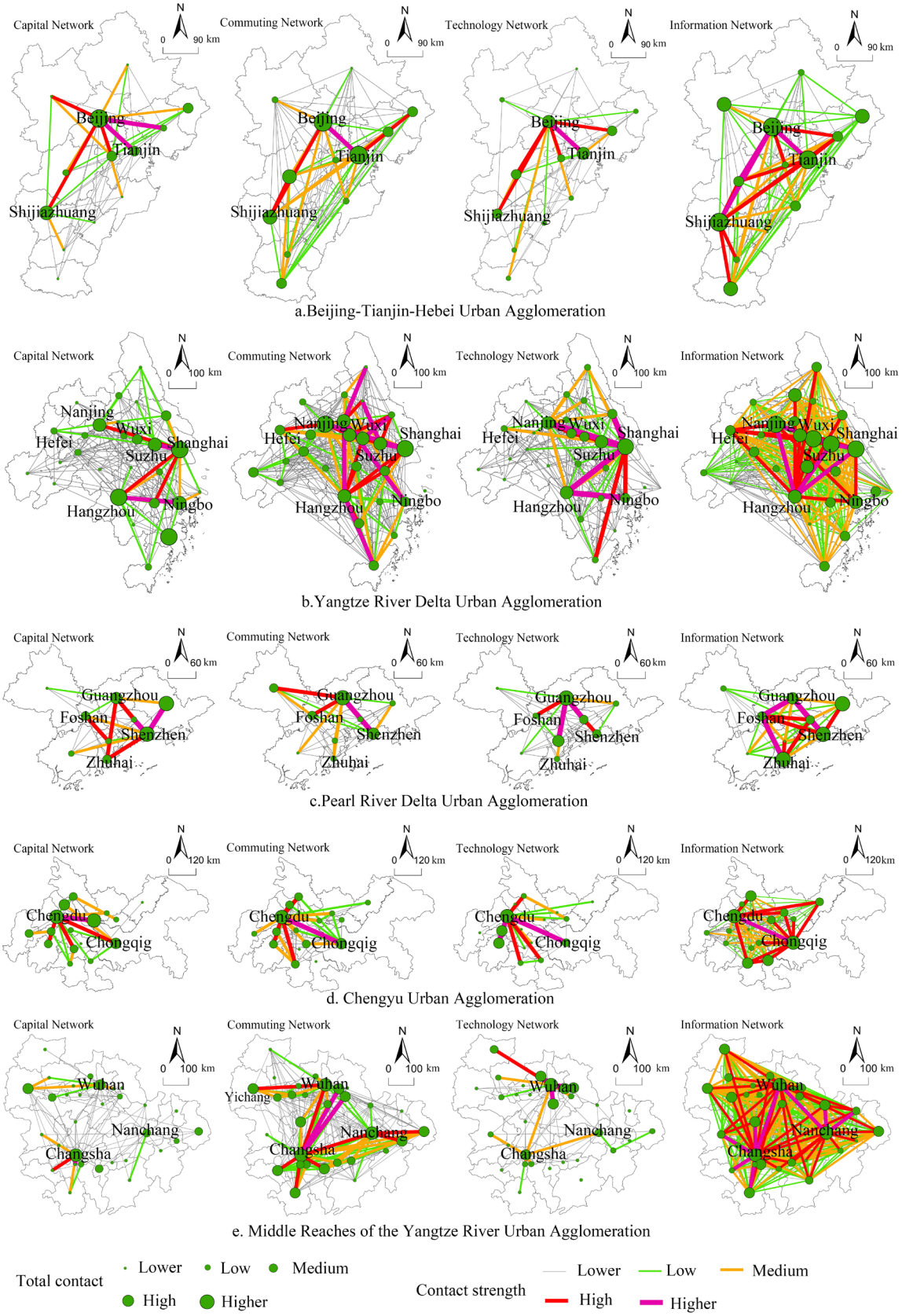
4.1.1 City node level

Beijing, Shanghai and Chengdu have always been the most connected cities in their respective urban agglomerations. The Pearl River Delta shows a dual-core competition model between Guangzhou and Shenzhen, and the urban agglomeration in the middle reaches of the Yangtze River presents three core patterns. In the Beijing–Tianjin–Hebei urban agglomeration, Tianjin is at the first level of commuting and information connections, Shijiazhuang is at the first level of information connection, and the total number of connections in other cities is relatively low. This is related to the long-standing prominent dualistic structure of regional development in the Beijing–Tianjin–Hebei urban agglomeration, with Beijing and Tianjin having a better foundation for development and other cities having a larger development gap with these two cities. In the Yangtze River Delta city cluster, Shanghai and Hangzhou are at the first level of capital connection, ahead of Suzhou, Nanjing and Ningbo. Shanghai and Nanjing are at the first level of commuting connections, whereas cities such as Suzhou, Wuxi and Hangzhou are at the second level. Shanghai is dominant in technological connections, while Suzhou, Nanjing and Hangzhou are at the second level. The information connections show the multihead patterns of Shanghai, Suzhou, Wuxi, Nanjing and Hangzhou. The spatial pattern of these network nodes is consistent with the characteristics of the polycentric structure of the Yangtze River Delta. Shenzhen and Huizhou have the highest total capital connections among the Pearl River Delta urban agglomerations, while Guangzhou ranks second. This shows that in terms of industrial spillover and foreign investment in Shenzhen, Huizhou has become the city of choice, and the external radiation effect of Guangzhou's capital is still insufficient. Guangzhou is a railroad hub city with significantly more commuting links than Shenzhen, and Zhaoqing is at the third level. Guangzhou has the largest number of scientific and technological connections, which are closely related to the concentration of its high-level research institutes and universities. Guangzhou’s scale of foreign scientific and technological cooperation is significantly greater than that in Shenzhen and other cities. Zhongshan, on the other hand, lacks the resources of scientific research institutes and universities and relies mainly on external cooperation for innovation; thus, its scale of external scientific and technological cooperation is large. Shenzhen, Zhuhai and Huizhou rank first in information connections, whereas Guangzhou ranks second. In the Chengdu–Chongqing urban agglomeration, the total number of connections in Chengdu is generally greater than in Chongqing, which is related to the administrative boundary barrier. The number of multidimensional connections in Chengdu is at the highest level. Chongqing's information connections are at the first level, commuting and capital connections are at the second level, and scientific and technological connections are at the third level. This is related to the fact that 14 of the Chengdu–Chongqing urban agglomerations are located in Sichuan Province and have close interconnections with each other yet relatively weak connections with Chongqing. In addition, the nodes are tripartite in the urban agglomeration of the middle reaches of the Yangtze River. Wuhan's total number of connections ranks first, except for capital connections. Changsha ranks first in technological connections and other connections. The total number of capital and technological connections in Nanchang is relatively weak.

4.1.2 Urban core axis

On the one hand, from the point of view of network connection strength and node status, network connection strength shows a significant geographical attenuation law, and the characteristics of the core edge are more significant. On the other hand, from the point of view of the node form of the network connection, the network connection presents a "single center", "dual center" and "multicenter" network spatial relationship.

Figure 3. The multidimensional geographic network of urban agglomerations in China



The central cities of the five major urban agglomerations occupy the core position in the network connection, and the number of connections is relatively large, as shown in Fig. 3. In the Beijing–Tianjin–Hebei urban agglomeration, Beijing and Tianjin have the highest contact strengths. The economic connection framework of the Chengdu–Chongqing urban agglomeration revolves around the central cities of Chengdu and Chongqing. The urban agglomeration in the middle reaches of the Yangtze River is centered on three core cities: Wuhan, Changsha and Nanchang. The Yangtze River Delta urban agglomeration forms a core framework around a few central cities, including Shanghai, Hangzhou, Suzhou and Nanjing.

4.2 Overall Network Structure Comparison

4.2.1 Overall network density characteristics

The density of the economic connection network is between 0.1 and 0.3, indicating that the urban agglomeration has spatial correlation characteristics that form a multithread, multicity and cross-regional network distribution trend. The economic connection network density of the Chengdu–Chongqing urban agglomerations exceeds 0.267, indicating they have a small network development gap, relatively balanced internal development, and relatively close connections. The network density of the Pearl River Delta urban agglomeration is only 0.111. There is a large gap in the development of the internal network of the Pearl River Delta urban agglomeration, which is more concentrated in the connections between a few internal core cities, and the siphoning effect of the central cities is more significant. The network densities of Beijing–Tianjin–Hebei, the Yangtze River Delta and the middle reaches of the Yangtze River are greater than those of the Pearl River Delta, indicating a relatively high network coordination of the three regions and a relatively small development gap within the network.

4.2.2 Stability characteristics of the spatial network

In addition to some indicators, the multidimensional network efficiency of the five urban agglomerations exceeded 0.6, indicating that more than 60% of the connections are effective. However, there are a few repeated superposition phenomena of dynamic correlation between urban agglomeration networks, indicating that the stability of the multidimensional network of the five major urban agglomerations is still insufficient. The network efficiency of technology and capital connections is relatively high, indicating that the superposition effect of the science and technology and capital connection networks in the five major urban agglomerations is insufficient. This means there are more independent connection nodes in the networks in these urban agglomerations and they have significant network edge areas.

4.2.3 Correlation characteristics of the spatial network

Social network analysis usually uses "average distance" to quantitatively reveal the small-world characteristics of the network (Ducruet & Beauguitte, 2014). The average distance of the economic, capital, commuting, technology and information connections of the five major urban agglomerations in China is between 1 and 3. The largest average distance of the capital connection networks of the urban agglomerations in the middle reaches of the Yangtze River is only 2.488, as shown in Table 1. Any two urban nodes in different types of spatial networks of the five urban agglomerations could establish connections with 1–3 intermediate cities; that is, the urban connection network presents significant small-world characteristics. The characteristics of spatial association networks promote interactions between urban connection networks, and the superposition effect of multidimensional urban networks on urban economic connections is prominent.

Table 1 Network structure characteristics

Network indicators	Contact Network	Beijing-Tianjin-Hebei	Yangtze River Delta	Pearl River Delta	Chengdu-Chongqing	Yangtze river middle reaches
Density	Capital	0.128	0.111	0.208	0.113	0.072
	Commuting	0.269	0.272	0.306	0.254	0.209
	Technology	0.179	0.119	0.833	0.154	0.084
	Information	0.339	0.291	0.333	0.254	0.358
	Economic	0.179	0.188	0.111	0.267	0.148
Efficiency	Capital	0.909	0.858	0.857	0.897	0.953
	Commuting	0.756	0.697	0.893	0.756	0.803
	Technology	0.946	0.913	0.214	0.956	0.946
	Information	0.636	0.643	0.643	0.743	0.621

Network indicators	Contact Network	Beijing-Tianjin-Hebei	Yangtze River Delta	Pearl River Delta	Chengdu-Chongqing	Yangtze river middle reaches
	Economic	0.861	0.809	0.667	0.838	0.880
Average distance	Capital	2.012	1.111	1.778	2.140	2.448
	Commuting	1.727	1.917	1.833	1.665	1.967
	Technology	1.788	2.071	1.167	1.829	2.347
	Information	1.509	1.591	1.400	1.729	1.642
	Economic	2.067	2.079	1.333	1.767	2.151

5. Multidimensional Network Correlation Effect of Urban Agglomerations

5.1 Correlation Analysis

In the five urban agglomerations, all correlation coefficients between multidimensional networks are positive. There is a significant positive correlation between the economic, capital, commuting, technology, and information connections of the five urban agglomerations in China.

On the one hand, from a dimensional perspective, the correlations between the information and economic connections of urban agglomerations in the Pearl River Delta are significant at the 5% level, and those between the information and economic connections of urban agglomerations in Beijing–Tianjin–Hebei, the Yangtze River Delta, Chengdu–Chongqing and the middle reaches of the Yangtze River are significant at the 1% level. The overall correlation coefficient remains between 0.4 and 0.7, the value of which is relatively average. The correlation between commuter and economic connections passed the 5% significance level test, and those between the urban agglomerations of Beijing–Tianjin–Hebei, the Yangtze River Delta, Chengdu–Chongqing and the middle reaches of the Yangtze River passed the 1% significance level test. The overall correlation coefficient is 0.24–0.72. In addition to the Pearl River Delta urban agglomeration, the correlations between technological and economic connections in Beijing–Tianjin–Hebei, the Yangtze River Delta, Chengdu–Chongqing and the middle reaches of the Yangtze River urban agglomeration passed the significance test at the 1% level, and the correlation coefficients ranged from 0.3 to 0.94. The spatial correlation coefficient between capital and economic connections is significantly lower than that between information, commuting and technological connections. These results show that information, commuting and technology connections greatly impact regional economic connections.

On the other hand, from a regional perspective, the spatial correlation coefficients between economic connections and the capital, commuting, technology and information connection indicators are relatively high in the Beijing–Tianjin–Hebei urban agglomeration. The correlation coefficient of technological connections reached 0.94, indicating the Beijing–Tianjin–Hebei regional connections are more dependent on scientific and technological cooperation. The degree of economic connection of the Yangtze River Delta shows great volatility, with a correlation coefficient of only 0.143, indicating a low impact of capital connections on the Yangtze River Delta's economic connection network, which depends more on commuting and information connections. The Pearl River Delta urban agglomeration is consistent with the Yangtze River Delta urban agglomeration, and commuting contact and information contact play a significant role in economic relations. The role of the science and technology connection network of the Chengdu–Chongqing urban agglomeration lags behind and depends more on the capital and commuter connection networks. The urban agglomeration in the middle reaches of the Yangtze River is consistent with the Beijing–Tianjin–Hebei urban agglomeration and relies more on patent cooperation; however, the overall correlation is not high.

Table 2 QAP correlation analysis of the spatial correlation of the economy

Variable	Beijing–Tianjin–Hebei	Yangtze River Delta	Pearl River Delta	Chengdu–Chongqing	The Yangtze river middle reaches
Capital	0.642***	0.143***	0.05	0.677***	0.271***
Commuting	0.724***	0.558***	0.452**	0.689***	0.243***
Technology	0.940***	0.452***	0.306*	0.449***	0.517***
Information	0.696***	0.518***	0.408**	0.584***	0.420***

Note: *, * *, * * * represent the significance levels of 10%, 5% and 1%, respectively.

5.2 Network Regression Analysis

On the basis of the correlation analysis, QAP regression analysis of the spatial correlation of the economic urban network is conducted to further explore the spatial connection and role of the information, commuter, technology and capital connections in economic connections, as shown in Table 3. Although the adjusted R2 of the Pearl River Delta urban agglomeration passed the significance test at the 5% level, the R2 values of the four urban agglomerations of Beijing–Tianjin–Hebei, the Pearl River Delta, Chengdu–Chongqing and the middle reaches of the Yangtze River all passed the significance test at the 1% level. According to the numerical analysis, the R2 of the Beijing–Tianjin–Hebei urban agglomeration is 0.882, which is the highest. The explanatory capacities of the Yangtze River Delta, the middle reaches of the Yangtze River and the Chengdu–Chongqing urban agglomerations are 63.2%, 53.6% and 48.3%, respectively. However, the explanatory capacity of the Pearl River Delta is low, at only 23.7%, which shows that capital, commuting, technology and information connections have a good explanatory capacity for the economic connections in the four urban agglomerations of Beijing–Tianjin–Hebei, the Pearl River Delta, Chengdu–Chongqing and the middle reaches of the Yangtze River.

In terms of dimensions, information connections passed the 10% significance level test, and three urban agglomerations passed the 1% significance test. The regression coefficients are relatively average and positive. Except for the urban agglomerations in the middle reaches of the Yangtze River Basin, the commuting connections of other urban agglomerations passed the 10% significance test, but the regression coefficient had two negative values, which remained in the range of [-0.025,0.396]. This finding indicates that the network relationships represented by the HSR train frequency data are not characterized by a regular distance decay pattern but rather by a "leapfrog" linking relationship with strong connections, mainly between large cities. Three urban agglomerations of technological linkages passed the 10% significance test, but the coefficient fluctuation was the largest, in the range of [-0.047,0.970]. Only two urban agglomerations passed the 5% significance test, and one urban agglomeration passed the 1% significance test for capital connections. Therefore, the subindicators of capital, commuting, technology, and information connections play an average role in supporting overall economic connections.

In terms of the regions, the regression coefficient of patent cooperation in the Beijing–Tianjin–Hebei urban agglomeration is 0.97, which passes the 1% significance test, whereas the regression coefficients of the other indicators are small, indicating that technological connections play a very significant role in supporting economic connections. The high-speed train frequency and Baidu index of the Yangtze River Delta urban agglomeration passed the significance test at the 1% significance level, and the regression coefficients were 0.396 and 0.222, respectively, indicating that the economic connection of the Yangtze River Delta urban agglomeration depends more on information and commuter connections. The Baidu index of the Pearl River Delta urban agglomeration passed the significance test at the 1% level, and the coefficient is the highest, indicating the economic connection network depends more on information connections. The regression coefficient of the frequency of high-speed trains in the Chengdu–Chongqing urban agglomeration is the highest at the 1% significance level, indicating that the economic connection network depends more on the commuter connection network. The patent cooperation, Baidu Index and capital connection of the urban agglomeration in the middle reaches of the Yangtze River passed the 1% significance test, reaching 0.405, indicating that economic connections depend more on the patent network.

Table 3 QAP regression analysis of the spatial correlation of the economy

Variable	Beijing-Tian-jin-Hebei	Yangtze Riv-er Delta	Pearl River Delta	Cheng-du-Chongqing	The Yangtze river middle reaches
Intercept	47.470	-50.578	-305.812	6.171	0.335
Capital	0.005	-0.007	-0.191**	0.051	0.140***
Commuting	-0.025**	0.396***	0.263*	0.159***	-0.023
Technology	0.970***	0.085	0.126*	-0.047	0.405***
Information	0.020*	0.222***	0.296***	0.156**	0.228***
Adj-R2	0.882***	0.632***	0.237**	0.483***	0.536***
Obs	156	702	72	240	930

Note: *, * *, * * * respectively represent the significance levels of 10%, 5% and 1%.

6. Conclusions and Implications

6.1 Research Conclusions

According to the general theory of economic growth and multidimensional flow data, a theoretical model of multidimensional network connections in urban agglomerations is constructed, and an empirical analysis is conducted on the five major urban agglomerations in China. The main conclusions are as follows: (1) The multidimensional connection network of urban agglomerations presents significant spatial differences in the "point–line-network", and the heterogeneity of the spatial structure of multidimensional connection networks in different dimensions and urban agglomerations is highly significant. (2) The multidimensional spatial correlation network of urban agglomerations presents multithreaded, multicity, and cross-regional network distribution trends, with significant "local area network" and "small world" characteristics. (3) The QAP regression analysis reveals that the capital, commuting, technological and information connections of the five major urban agglomerations have a significant effect on economic connections. The impact of different dimensions of connections between different cities on economic connections is heterogeneous, and the four-dimensional connection network of Beijing–Tianjin–Hebei has a greater supporting effect on economic connections. The effects of different indicators on the economic connection network are similar, but the influence varies greatly among different urban agglomerations. The urban agglomerations of Beijing–Tianjin–Hebei and the middle reaches of the Yangtze River depend more on technology connections, the Yangtze River Delta urban agglomeration rely more on commuting connections, and the Pearl River Delta and Chengdu–Chongqing urban agglomerations rely more on information connections.

6.2 Policy Implications

First, the coordinated development of multidimensional connections among urban agglomerations should be promoted. An urban connection is a multidimensional comprehensive complex network system in which connection networks of different dimensions impact one another through economic connections. Therefore, attention should be given to promoting the coordinated development of capital, commuting, technological, and information connections in urban agglomerations. On the one hand, it is necessary to improve the hard connections within urban agglomerations. The construction of transportation, energy, information, financial networks and other facilities in urban agglomerations should be accelerated, and the carriers of factor flow should be optimized and improved. In particular, rapid transportation's driving role should be considered to accelerate the flow efficiency of human, capital, information, and technology flows. On the other hand, the soft connections within urban agglomerations should be optimized. The policy system for constructing urban agglomerations could be improved to break down internal administrative barriers. Furthermore, the integration of internal registered residences, medical care, education and other systems is explored to optimize the business environment, strengthen the stickiness of capital, population, technology, and information flows between cities, and enhance the spillover effect of central cities on small and medium-sized cities.

Second, urban agglomerations should be encouraged to develop their own strengths in multidimensional connections. Owing to differences in the resource environment, geographical location, industrial development and policy, the development of urban agglomerations in China are very heterogeneous, with differences in the supporting role of economic connections in the urban spatial network structures. The development of urban agglomerations in China remains unbalanced and insufficient, with significant internal development gaps. However, the differences in urban agglomeration network connections should be taken seriously, allowing for asymmetric development within a certain range and expanding and strengthening the core growth network. For example, the Beijing–Tianjin–Hebei region and the middle reaches of the Yangtze River urban agglomeration should leverage the advantages of patent cooperation between cities, further enhancing the overall level of technological network cooperation in the urban agglomeration. An urban agglomeration supported by science and technology networks should be built to promote the improvement of urban connections with scientific and technological support. For example, the Yangtze River Delta urban agglomeration should leverage the advantages of a well-developed high-speed rail network to further enhance the transportation accessibility of the urban agglomeration and create a rapid transportation-supported urban network.

Third, the connections between central cities and small and medium-sized cities should be strengthened. Small and medium-

sized cities should leverage the high-speed rail economy, knowledge economy, and information revolution; accelerate the connections of capital, technology, commuting, and information with central cities; and improve the economic connections between small and medium-sized cities and large cities. For example, in the context of the exchange opportunities created by high-speed rail, more opportunities for technology and industrial transfer should be provided, enhancing technology and commuting connections with central cities. In addition, central cities should expand their radiation and driving role in small and medium-sized cities and jointly build intercity collaborative innovation circles, industrial transfer parks, rapid commuting circles, and information-sharing platforms through targeted assistance and industrial transfers. The flow of elements between cities and the internal connections within urban agglomerations are increasingly promoted through agglomeration and diffusion.

6.3 Discussions

With the advent of globalization and informatization, factor flows have increasingly become an important force in shaping the spatial structure of regional economies, and "flow space" has become a hot issue in urban networks. The portrayal of the spatial structure of urban networks by a single flow element is insufficient, while multidimensional flow elements capture the spatial structure of urban agglomerations in a relatively comprehensive and adequate way. Since the network spatial structure of cities in different dimensions and different urban agglomerations is characterized by heterogeneity, the multidimensional flow spatial structure of multiple urban agglomerations is compared and analyzed, which has important theoretical and practical significance. This shows that the status of a city is not completely consistent with its economic scale, administrative level, or geographical location. It is necessary to pay attention to the high-level agglomeration of urban agglomeration network connections and to the balanced development of gradient differences, which are highly important for revealing the spatial structure of urban agglomeration networks more comprehensively.

Cities in the mobility space environment could transcend the limitations of geographic proximity to access key resources from distant partners; thus, the status of networks in the mobility space is becoming increasingly important in comparison to geographic location in the place-space environment. Factor mobility is increasingly embedded in different types of network connections; thus, the synergistic effects of multidimensional networks of factors of mobility should be considered. This indicates that the development of the urban economy is promoted by various factors breaking through local and regional spatial scales, and various factors are actively integrated into the domestic cycle and the national urban network, which provides a broader space for improving urban economic growth.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Aghion, P., & Howitt, P. (1998). Endogenous growth theory. The MIT Press.

[2] Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. American Economic Review, 86, 630–640.

[3] Barro, R., & Sala-i-Martin, X. (1992). Regional growth and migration: A Japan-United States comparison. Journal of the Japanese and International Economies, 6(b), 312–346.

[4] Bathelt, H., & Glückler, J. (2003). Toward a relational economic geography. Journal of Economic Geography, 3(2), 117–144. https://doi.org/10.1093/jeg/3.2.117

[5] Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. Progress in Human Geography, 28(1), 31–56.

[6] Beckman, C. M., Eisenhardt, K., & Kotha, S. (2012). The role of the entrepreneur in technology entrepreneurship. Strategic Entrepreneurship Journal, 6(3), 203–206.

[7] Bing, X., Xiao, X., & Jing, Z. (2020). Identification method and empirical study of urban industrial spatial relationship

based on POI big data: A case of Shenyang City, China. Geography and Sustainability, 1(2), 152–162.

[8] Brown, E., Derudder, B., & Parnreiter, C. (2010). World city networks and global commodity chains: Towards a world-systems integration. Global Networks, 10(1), 12–34.

[9] Burger, M. J., van der Knaap, B., & Wall, R. S. (2013). Polycentricity and the multiplexity of urban networks. European Planning Studies, 22(4), 816–840.

[10] Camagni, R., & Salone, C. (1993). Network urban structures in Northern Italy: Elements for a theoretical framework. Urban Studies, 30(6), 1053–1064.

[11] Cao, W., Du, D., & Xia, Q. (2023). Unbalanced global vaccine product trade pattern: A network perspective. Social Science & Medicine, 325, 115913.

[12] Castells, M. (1996). The rise of network society. Blackwell.

[13] Davoudi, S. (2008). Conceptions of the city-region: A critical review. Urban Design and Planning, 16(2), 51–60.

[14] Derudder, B., Taylor, P. J., & Ni, P. (2010). Pathways of change: Shifting connectivities in the world city network. Urban Studies, 47(9), 1861–1877.

[15] Ducruet, C., & Beauguitte, L. (2014). Spatial science and network science: Review and outcomes of a complex relationship. Networks and Spatial Economics, 14(3–4), 297–316.

[16] Gui, Q. C., Liu, C. L., & Du, D. B. (2019). Globalization of science and international scientific collaboration: A network perspective. Geoforum, 105, 1–12.

[17] He, C. F., Wei, Y. H. D., & Xie, X. Z. (2008). Globalization, institutional change, and industrial location: Economic transition and industrial concentration in China. Regional Studies, 42(7), 923–945. https://doi.org/10.1080/00343400701543272

[18] Hu, X., Wang, C., Wu, J., & Stanley, H. E. (2020). Understanding interurban networks from a multiplexity perspective. Cities, 99, 102625.

[19] Huber, F. (2012). Do clusters really matter for innovation practices in Information Technology? Questioning the significance of technological knowledge spillovers. Journal of Economic Geography, 12, 107–126.

[20] Huggins, R., & Prokop, D. (2017). Network structure and regional innovation: A study of university-industry ties. Urban Studies, 54(4), 931–952.

[21] Huggins, R., & Thompson, P. (2013). A network-based view of regional growth. Journal of Economic Geography, 14(3), 511–545.

[22] Huggins, R., & Johnston, A. (2010). Knowledge flow and inter-firm networks: The influence of network resources, spatial proximity, and firm size. Entrepreneurship and Regional Development, 22, 457–484.

[23] Jacobs, W., Koster, H., & Hall, P. (2011). The location and global network structure of maritime advanced producer services. Urban Studies, 48(13), 2749–2769.

[24] Jiao, J. J., Zhang, Q. L., & Wu, Y. Y. (2021). Change and influencing factors of China's cross-regional investment network structure. Progress in Geography, 40(8), 1257–1268.

[25] Jung, W. S., Wang, F., & Stanley, H. E. (2008). Gravity model in the Korean highway. EPL, 81(4), 48005.

[26] Justin, Y. L., Wang, G. W., & Zhao, Y. H. (2004). Regional inequality and labor transfers in China. Economic Development and Cultural Change, 52(3), 587–603.

[27] Kramera, J.-P., Marinelli, E., Iammarino, S., & Revilla Diez, J. (2011). Intangible assets as drivers of innovation: Empirical evidence on multinational enterprises in German and UK regional systems of innovation. Technovation, 31, 447–458.

[28] Li, P. F. (2014). Global temporary networks of clusters: Structures and dynamics of trade fairs in Asian economies. Journal of Economic Geography, 5, 995–1021.

[29] Liu, X., Derudder, B., & Garcia, C. G. (2013). Exploring the co-evolution of the geographies of air transport aviation and corporate networks. Journal of Transport Geography, 30(3), 26–36.

[30] Liu, Y., Kang, C. G., & Gao, S. (2012). Understanding intra-urban trip patterns from taxi trajectory data. Journal of

Geographical Systems, 14(4), 463–483.

[31] Lu, D. B., Xiao, W., Xu, G. Y., Ha, L., & Yang, D. Y. (2021). Spatiotemporal patterns and influencing factors of human migration networks in China during COVID-19. Geography and Sustainability, 2, 264–274.

[32] Lucas, R. E. (1988). On the mechanics of economic development. Journal of Monetary Economics, 22, 3–32.

[33] Ma, H., Wei, Y., & Huang, X. (2023). The innovation networks shaped by large innovative enterprises in urban China. Journal of Geographical Sciences, 33, 599–617.

[34] Malecki, E. J. (2002a). Hard and soft networks for urban competitiveness. Urban Studies, 39(5), 929–945.

[35] Malecki, E. J. (2002b). The economic geography of the Internet's infrastructure. Economic Geography, 78(4), 399–424.

[36] Marshall, A. (1920). Principles of economics. Macmillan.

[37] Peter, J., Taylor, M. H., & Verbruggen, R. (2010). External urban relational process: Introducing central flow theory to complement central place theory. Urban Studies, 47(13), 2803–2818.

[38] Pflieger, G., & Rozenblat, C. (2010). Introduction. Urban networks and network theory: The city as the connector of multiple networks. Urban Studies, 47(13), 2723–2735.

[39] Pflieger, G., & Rozenblat, C. (2010). Urban networks and network theory: The city as the connector of multiple networks. Urban Studies, 47(13), 2723–2735.

[40] Qian, X. Y., & Sun, B. D. (2021). The spatial structures and organization patterns of China's city network based on inter-city startup investments. Geographical Research, 40(2), 419–430.

[41] Romer, P. M. (1986). Increasing returns and long-run growth. Journal of Political Economy, 94(5), 1002–1037.

[42] Storper, M., & Venables, A. J. (2004). Buzz: Face-to-face contact and the urban economy. Journal of Economic Geography, 4, 351–370.

[43] Taylor, A. M., & Williamson, J. G. (2006). Convergence in the age of mass migration. European Review of Economic History, 1(1), 27–63.

[44] Tenkanen, H., Pereira, R., Arcaute, E., & Gonzalez, M. (2023). Advances in geospatial approaches to transport networks and sustainable mobility. Environment and Planning B: Urban Analytics and City Science, 50(8), 2009–2016.

[45] Tirthankar, B., Arijit, D., & Paulo, P. (2023). Exploring the drivers of urban expansion in a medium-class urban agglomeration in India using the remote sensing techniques and geographically weighted models. Geography and Sustainability, 4(2), 150–160.

[46] Xu, J., Yang, Y., Zhang, Y., & Man, S. (2024). 'Partly' globalized networks and driving mechanism in resource-based state-owned enterprises: A case study of J Group. Geography and Sustainability, 5(1), 77–88.

[47] Xu, X., & Zhao, M. F. (2020). Data capital and economic growth path. Economic Research Journal, 55(10), 38–54.

[48] Yang, H., Dijst, M., & Witte, P. (2019). Comparing passenger flow and time schedule data to analyze High-Speed Railways and urban networks in China. Urban Studies, 56(6), 1267–1287.

[49] Yang, S., Li, Z., Zhou, J., Gao, Y., & Cui, X. (2024). Evolving patterns of agricultural production space in China: A network-based approach. Geography and Sustainability, 5(1), 121–134.

[50] Yao, S. M., Zhu, Y. M., & Chen, Z. G. (2001). Urban agglomerations in China. University of Science and Technology of China Press.

[51] Zhang, W. Y., Derudder, B., Wang, J. H., & Witlox, F. (2020). An analysis of the determinants of the multiplex urban networks in the Yangtze River Delta. Tijdschrift Voor Economische En Sociale Geografie, 111(2), 117–133. https://doi.org/10.1111/tesg.12361

[52] Zhang, Y., Zhang, A., & Wang, J. (2020). Exploring the roles of high-speed train, air and coach services in the spread of COVID-19 in China. Social Science Electronic Publishing.

[53] Zhen, F., Qin, X., & Ye, X. (2019). Analyzing urban development patterns based on the flow analysis method. Cities, 86, 178–197.

[54] Zhen, F., Wang, X., & Yun, J. (2013). An empirical study on Chinese city network pattern based on producer services. Chinese Geographical Science.

Asia Pacific Science Press

Asia Pacific Economic and Management Review
ISSN(O):3005-9275
ISSN(P):3005-9267
Vol. 3 No.3 (2026)

Research on Audit Agents Based on RPA Technology: A Case Study of Deloitte's Financial Robot

Yanan Wei*, Xuandong Zhang

School of Economics and Management, Yan'an University, Yan'an, Shaanxi, 716000, China

Corresponding author: Yanan Wei, yha@yau.edu.cn

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Traditional audit processes involve numerous repetitive, rule-based manual tasks that are inefficient and prone to errors. The emergence of RPA (Robotic Process Automation) technology offers a viable solution to these challenges. Using Deloitte's financial automation tool "Xiao Qin Ren" as a case study, this paper employs a single-case analysis approach to examine RPA's application scenarios, operational mechanisms, and implementation outcomes in auditing, aiming to derive practical insights and theoretical [a]implications through detailed examination. The study reveals that "Xiao Qin Ren" can automatically perform repetitive tasks such as account reconciliation, confirmation letter generation, and working paper preparation by simulating human operations. Its deployment reduced reconciliation time from 3 person-days to just 10 minutes with 100% accuracy. According to Deloitte's official case report, during a proof-of-concept (POC) at a bank, the average time required to obtain individual audit evidence decreased from 40 minutes to under 30 seconds. This case demonstrates the applicability of RPA technology in auditing and provides valuable insights for accounting firms pursuing intelligent audit transformation.

Keywords: RPA; Intelligent Audit; Financial Robotics; Audit Automation

Published: Jun 24, 2026

DOI: https://doi.org/10.62177/apemr.v3i3.1484

1. Introduction

1.1 Research Background

Amid the wave of digital transformation, efficiency bottlenecks in traditional audit workflows have become increasingly evident. Conventional auditing processes rely heavily on manually performed tasks with well-defined yet highly repetitive rules, which not only consume significant human resources but also risk compromising audit accuracy due to human error—making industry-wide transformation imperative[1]. The core challenge of current transformation technologies lies in achieving faster audits at lower costs while maintaining reliable quality. The maturation of RPA (Robotic Process Automation) technology offers a novel solution: as software that replaces humans in performing high-repetition tasks based on predefined rules, it seamlessly mimics human behavior to operate information systems and replicate standardized workflows, perfectly suited for auditing needs[2]. Notably, the audit automation tool "Xiaoqinren," jointly developed by Deloitte and KiraSystems

a. A "3-person-day" is a unit of workload, which simply means one person working for three days, or three people working for one day—totaling three "person-days" of effort.

in 2016, has demonstrated proven effectiveness in real-world applications, providing substantial practical value for related research.

1.2 Research Objectives

Deloitte, as one of the Big Four accounting firms, is widely recognized as a benchmark in the auditing industry. "Xiaoqinren," one of the earliest and most prominent RPA products in this field, has undergone years of implementation and iteration, achieving highly mature technical capabilities supported by extensive publicly available research and practical documentation. To date, the product has provided process automation services to over 200 organizations, with proven application maturity and traceable, verifiable implementation outcomes. Deloitte's strategic approach of "technology-driven business" serves as a practical model for domestic peers and offers significant replicability and reference value for the widespread adoption of RPA solutions across the industry[3].

The purpose of this study is to address three core questions through the Deloitte "Little Deloitte People" case study: First, what audit challenges can RPA solve? Second, how exactly does it achieve these solutions? Third, what are the practical outcomes? The research aims to theoretically validate the applicability of RPA technology in auditing scenarios and provide actionable guidance for the intelligent transformation of the auditing industry.

1.3 Research Significance

Regulatory requirements for audit practice quality continue to tighten, while clients' expectations regarding audit timeliness and value-added services are steadily rising. The traditional audit model can no longer meet these dual demands. The Big Four accounting firms have pioneered the use of RPA technology to streamline workflows, establishing a significant competitive advantage[4]. However, most domestic firms still rely primarily on conventional manual processes, which not only fail to meet regulatory quality verification requirements but also cause them to fall behind in market competition. Digital transformation has thus become an urgent necessity.

2. Case Background

Deloitte is a globally leading professional services firm whose business encompasses auditing and certification, tax and legal services, consulting, and financial advisory. It serves a broad and high-quality client base. Amid the digital transformation wave, Deloitte pioneered the "Smart Audit" transformation strategy, positioning technological innovation as the core driver for enhancing service competitiveness and continuously exploring ways to apply cutting-edge technologies in professional service scenarios, establishing itself as a benchmark enterprise for industry digital transformation[5].

2.1 Overview of Pain Points

2.1.1 Pain Points in the Audit Industry

The traditional audit model has long been plagued by common challenges across four key dimensions—efficiency, cost, quality, and duration—as detailed in Table 1 below:

Table 1 Common Pain Points of Traditional Audit Models

Pain Point Dimension	embody	Data Support
productiveness	Repetitive tasks such as account reconciliation and confirmation requests consume a significant amount of time.	The average monthly time required for account reconciliation at a medium-sized company is 3 person-days.
prime cost	Professionals are engaged in jobs with low added value.	Human costs account for over 60% of the firm's operating expenses.
quality	Manual operation is prone to errors.	The manual error rate in the audit process is approximately 1-3%.
period	The audit process is lengthy, and customers have to wait a long time.	The annual audit of listed companies typically takes 1 to 3 months.

These pain points not only hinder the improvement of audit industry efficiency but also make it difficult to further enhance customer satisfaction, necessitating urgent technological solutions to overcome these development bottlenecks.

2.1.2 The challenges Deloitte once faced

As the market environment evolves, Deloitte's traditional labor-intensive audit model is increasingly proving unsustainable: On one hand, clients are demanding higher audit efficiency—most enterprises expect audit cycles to be reduced by over 30% compared to previous levels, while also requiring audit services to deliver higher-value outputs such as risk assessment and process optimization; on the other hand, labor costs continue to rise annually, with a significant number of certified public accountants with professional expertise being occupied by repetitive tasks, leading to a sustained decline in project profit margins. How to reduce operational costs and enhance service efficiency while strictly maintaining audit quality has become a core challenge for Deloitte, making the transition to intelligent auditing imperative[6].

2.2 Case Theme and Decision-Making Process

2.2.1 Why choose RPA technology?

During the technology selection phase, Deloitte conducted a comparative evaluation of various digital solutions: General AI large models feature complex technical logic and high training costs, while audit scenarios demand exceptionally high result accuracy, making complete reliance on AI's autonomous judgment challenging; traditional customized software lacks sufficient flexibility, requires integration with multiple heterogeneous systems within clients' and audit firms' environments, entails high modification costs and lengthy implementation cycles, and struggles to meet diverse client-specific needs[7]. In contrast, RPA (Robotic Process Automation) offers significant advantages: as a non-invasive technology, it requires no system modifications, performs tasks by simulating human operation interfaces, and can be deployed based solely on predefined rules, featuring low implementation difficulty and rapid deliverability—making it perfectly suited for the numerous standardized, repetitive tasks inherent in audit processes and the ideal choice for core technological applications.

2.2.2 Deloitte's RPA Strategic Framework

Following its official launch of RPA application exploration in the audit field, Deloitte assembled a dedicated RPA team comprising audit experts and technical developers to systematically analyze standardized scenarios across the entire audit process. After multiple rounds of feasibility assessments, Deloitte opted against purchasing off-the-shelf solutions and instead developed a proprietary RPA product leveraging its deep expertise in audit operations, establishing a distinctive technological advantage under the brand name "Xiao Qin Ren".

2.3 Actions Taken

2.3.1 Development Timeline for "Little Diligent Person"

The implementation of the "Little Diligent Person" initiative proceeds steadily in four phases, with specific timelines and tasks detailed in Table 2 below:

Table 2: The four stages of implementation for the "Little Diligent Person" initiative

stage	time	ground work
Exploration Phase	Around 2016	Conduct a comprehensive audit of standardized processes throughout the entire investigation and auditing workflow to evaluate the feasibility of implementing RPA technology.
development period	2016–2017	completed the development of core functional modules and launched the first prototype of "Xiaoqinren"
Pilot Period	2017	Select 12 audit projects from the manufacturing and financial sectors for pilot implementation to continuously optimize product logic.
Promotion Period	Since 2018	Deploy and promote globally, continuously iterate features and release updates based on business needs

2.3.2 Explanation of the Technical Principles of "Xiao Qin Ren"

"Xiaoqinren" is essentially a computer programming software robot specifically designed for audit scenarios, primarily simulating human computer operations: using predefined rule scripts, it automatically performs tasks such as mouse clicks, keyboard inputs, data reading, and format validation without requiring integration with system-level interfaces, and runs directly on the user interfaces of existing online banking systems, audit systems, and office software. With clearly defined task rules and standardized operational steps, corresponding automated workflows can be quickly configured without complex secondary development.

2.3.3 Main Functional Modules of "Little Diligent Person"

"Xiaoqinren" has developed four standardized functional modules centered around the core audit processes:

Function 1: Automatic Statement Generation

You can download bank statement data in bulk, automatically compare each entry with corporate account records, and quickly identify abnormal discrepancies.

Function 2: Automatic Generation of Confirmation Letters

Directly extract customer information, account details, and other data from the audit system, generate standardized confirmation letters in batches, and automatically send them to the designated email addresses of the respective parties being confirmed.

Function 3: Auto-populate Audit Working Papers

The reconciliation results, confirmation response details, third-party data, and other content will be automatically populated into the corresponding sections of the audit working papers without manual copying or pasting.

Function 4: Auto-generate variance report

Summarize all anomalies across the entire process, automatically categorize and organize them to generate a standardized variance report for direct review by auditors.

2.4 Actual Situation and Outcomes

2.4.1 Promotion and Application Status

By Dec 2025, "Xiaoqinren" had been deployed in over 100 Deloitte offices worldwide, serving clients across more than 20 industries including finance, manufacturing, retail, and high-tech sectors. It has been implemented in over 3,000 audit projects, becoming one of the core tools within Deloitte's Smart Audit System. Its adoption model has accelerated the widespread adoption of RPA solutions across the audit industry, with performance data for typical use cases presented in Table 3.

Table 3 Effects of Typical Application Scenarios

application scenarios	Traditional Method	After using Xiaoqinren:	Efficiency Improvement
Bank Statement (Monthly)	3 person-days	10 minutes	Approximately 95%
confirmation letter generation (100 copies)	1 person per day	15 minutes	Approximately 95%
Draft Completion (Single Subject)	2 hours	5 minutes	Approximately 95%
continuous work	Work 8 hours daily and require rest.	operates continuously for 24 hours	/
precision	Approximately 97–99%	100% (Within Rules)	markedly improve

3. Case Analysis

3.1 Path Analysis: How Does RPA Achieve Audit Automation?

Table 4 below presents Xiaqinren's tailored solutions to the key challenges of traditional auditing:

Table 4 Solutions by "Xiaoqinren" for Traditional Audit Pain Points

Traditional Pain Points	Xiaoqin People Solution	Implementation Principle
Account reconciliation takes a long time	Auto-login, Download, Verify	Simulate Manual Operation + Rule Matching
The processing of confirmation requests is slow.	Batch generate and send automatically	Data Extraction+Template Population
The draft form is cumbersome to fill out.	Auto-fill	Data Mapping+Automatic Entry
Humans are prone to errors.	Machine Execution, Unified Rules	Execute precisely according to the preset logic
Cannot work continuously	24 hours uninterrupted	The software-controlled robot does not experience any sense of fatigue.

When designing audit robots, priority should be given to tasks with low complexity and minimal variation according to the process pyramid. For high-level business processes that are highly specialized, unstable, or lack clearly defined standards, manual operation remains the preferred approach.

3.2 Effect Analysis

3.2.1 Efficiency and Performance

The implementation of "Xiaoqinren" has achieved a significant leap in audit efficiency. Taking bank reconciliation as an example, the time required to generate confirmation letters was reduced from one person-day to just 15 minutes, while the preparation time for working papers was cut from 1.5 hours to under 30 minutes—a 67% improvement in efficiency. In short, Deloitte leveraged "Xiaoqinren" to deliver a remarkable 40–60% increase in audit efficiency.

3.2.2 Quality Performance

Within predefined rules, RPA achieves 100% operational accuracy, eliminating verification errors caused by fatigue or negligence. Moreover, each robotic operation is digitally recorded, ensuring traceability and transparency in the audit process and meeting stringent internal control requirements.

3.2.3 Cost-Benefit Analysis

From an input-output perspective, the cost-effectiveness advantage of RPA systems is particularly pronounced: in routine audit projects, they can replace approximately 50–70% of repetitive tasks, enabling junior auditors responsible for such work to transition into more growth-oriented roles such as risk assessment and industry analysis; senior certified public accountants can focus more on professional judgment tasks, thereby effectively controlling overall labor costs. In terms of return on investment, the deployment cost of RPA for a single project is typically fully recovered through labor cost savings within 6–12 months, yielding significant long-term benefits.

4. Countermeasure Recommendations

4.1 Recommendations at the Technical Level

First, system compatibility must be enhanced by developing universal adaptation components for heterogeneous ERP systems, online banking platforms, and business systems across different industries, thereby reducing customization efforts and lowering technical barriers to cross-system deployment. Second, exception handling capabilities need improvement through optimized automatic identification logic for non-standard scenarios; when encountering rule-violating operational anomalies, the system should automatically pause operations and trigger alerts to prevent invalid executions or erroneous outputs. Addressing these challenges, Deloitte Xiaoqin has now evolved to its third generation, integrating IDRP intelligent document review, OCR (Optical Character Recognition), and NLP (Natural Language Processing) technologies in a deep integration. It not only handles data transfer but also understands content, with its core strength lying in processing unstructured data and supporting over 20 document formats[8]. Further deep integration will be pursued in the future.

4.2 Recommendations at the Promotion Level

On one hand, it is essential to lower the entry barrier for frontline staff; ordinary auditors can quickly establish automated processes through drag-and-drop operations and parameter configuration without requiring coding skills, thereby reducing reliance on technical teams. On the other hand, personnel skill training must be strengthened by incorporating RPA operations and process configuration into regular audit training programs, fostering versatile professionals who understand audit business logic while being proficient in digital tools. Concurrently, a cross-industry best practice repository should be established, compiling mature RPA application cases from clients across various sectors and sizes for project teams to reference and reuse, thereby minimizing trial-and-error costs. Deloitte conducted RPA training courses in Beijing, Hong Kong, and other locations in 2019, where 84 audit professionals with no IT background mastered relevant development skills after six days of training[9]. This approach warrants widespread adoption.

4.3 Recommendations at the Institutional Level

Industry regulators and associations can lead the development of unified standards for RPA audit applications, specifying common requirements such as technical security, operational protocols, and data access permissions to guide the industry-wide standardized use of digital tools. Within accounting firms, clear delineation of authority boundaries for human-machine collaboration is essential, along with well-defined scope of RPA application scenarios and accountability for reviewing machine-generated results, thereby mitigating audit risks arising from ambiguous responsibilities. A dynamic evaluation system should be established to regularly assess RPA implementation effectiveness across three dimensions—efficiency improvement, quality enhancement, and cost reduction—enabling continuous refinement of technical solutions and management mechanisms.

5. Conclusion and Prospects

5.1 Research Conclusion

This paper employs a single-case study approach to analyze the implementation of Deloitte's "Little Deloitte People," demonstrating the applicability and core value of RPA technology in the field of intelligent auditing. Leveraging its non-intrusive, rule-based characteristics, RPA effectively addresses efficiency challenges associated with repetitive tasks such as account reconciliation, confirmation requests, and working paper preparation in traditional auditing, achieving dual improvements in both efficiency and quality: efficiency for core standardized tasks increases by over 90%, operations within predefined rules are error-free, and it can replace more than 50% of low-value-added manual work. Furthermore, "Little Deloitte People" establishes a novel audit model based on human-machine collaboration, freeing auditors from mechanical tasks so they can focus on professional judgment and risk management, providing a practical example for the digital transformation of accounting firms.

5.2 Future Outlook

The future application of RPA in the auditing field will evolve along three key directions. First, technological integration: RPA will deeply integrate with AI and big data, evolving from process automation to intelligent decision support that enables risk identification and analysis in complex scenarios. Second, end-to-end application coverage: it will expand from single-process automation to encompass the entire audit lifecycle—including planning, execution, and review—establishing a comprehensive intelligent audit system. Third, widespread adoption: RPA will extend from leading institutions like the Big Four to domestic small and medium-sized accounting firms, as industry-wide standards and cost-effective tools continue to mature, driving digital and intelligent transformation across the auditing sector.

Funding

No

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Cheng, P., & Tang, C. R. (2025). Research on an intelligent enterprise information verification and audit agent based on RPA and DeepSeek: A case study of the ND Accounting Firm's project with the Market Supervision Bureau. Commercial

Accounting, (23), 4–9.

[2] Cheng, P., & Deng, J. H. (2023). Design and application of a tax refund invoice generation robot based on RPA technology. Finance and Accounting, (16), 54–57.

[3] Cheng, P., & Xiong, J. Y. (2023). Design and application of an intelligent expense reimbursement robot based on RPA + AI. Finance and Accounting, (10), 52–55.

[4] Guo, J., Cheng, P., Zhu, J. H., et al. (2026). Research on the pathways and effects of RPA in empowering financial digital transformation in water utilities. Accounting Monthly, 47(5), 85–91.

[5] Yun, L. X., Liu, Y. X., & Fan, Y. N. (2026). Research on the pathways of artificial intelligence in empowering enterprise intelligent finance development: A case study based on Deloitte. Journal of Hubei University of Economics (Humanities and Social Sciences Edition), 23(3), 44–50.

[6] Wu, Y., Li, Y. F., Huang, Y., et al. (2026). Human-machine collaborative intelligent audit system: Framework construction, decision-making paradigms and risk governance. China Certified Public Accountants, (1), 93–101+5.

[7] Sundarasen, S., Kamaludin, K., & Nakiran, D. (2026). From adoption to audit quality: Mapping the intellectual structure of artificial intelligence-enabled auditing. Journal of Risk and Financial Management, 19(3), 209.

[8] Zha, J. (2025). New quality productivity, optimization of internal audits in commercial banks, and intelligent collaboration. Finance Research Letters, 86, 108579.

[9] Qi, Y. (2025). The construction of a digital-intelligent audit theory system with Chinese characteristics in the new era under the digital context. Journal of Economic Research, 2(6), 9–11.

Dear Researchers and Scholars :

Greetings from Asia Pacific Science Press, a beacon of academic and scientific publishing, located in the vibrant city of Hong Kong.

We extend our heartfelt gratitude for your relentless pursuit of knowledge, and your significant contributions to the advancement of science and society. It is researchers and scholars like you who propel humanity forward, and we at the Asia Pacific Science Press are devoted to ensuring that your groundbreaking works receive the global recognition they rightfully deserve.

In light of our commitment to disseminating pioneering research across various disciplines, such as medicine, architecture, education, and electronics, we are reaching out with two pivotal opportunities to augment our collaboration with the global academic community:

Call for Paper Submissions:

We cordially invite you to submit your original research articles to our fast-growing, peer-reviewed, and open-access journals. Our platform guarantees an extensive, global reach, enabling your work to garner maximum visibility and citation in the academic sphere. Rest assured, your work will be meticulously assessed by experts in the field, ensuring it receives the acknowledgment and exposure it merits.

Join Our Esteemed Team:

We are fervently searching for passionate researchers and scholars interested in joining our burgeoning team at Asia Pacific Science Press. We offer numerous roles, such as peer reviewers, editors, and advisory board members, where your expertise will significantly shape the content and quality of our publications. In return, you will gain invaluable experience, network with preeminent scholars, and play a pivotal role in molding the future of global academic publishing.

Why Choose Asia Pacific Science Press?

Global Reach: Your work will be accessible to a worldwide audience, free from any access barriers.

Collaboration with Renowned Universities: We have established extensive publishing systems in cooperation with world-renowned universities, such as Wuhan University, Hong Kong University, and the University of Malaya.

Diverse Disciplines: Your research will be housed among numerous journals across a multitude of academic projects and disciplines.

As we stride forward in the academic landscape, we envision a future where our collective efforts shape a more enlightened, innovative, and interconnected global society. We sincerely hope that you consider this invitation to join us on this auspicious journey towards knowledge, discovery, and global impact.

Should you wish to submit your work or express interest in joining our team, please do not hesitate to contact us. You can submit your manuscript or personal profile to info@apspublisher.com or visit our website at www.apspublisher.com for more information.

Thank you for considering this opportunity, and we eagerly anticipate the possibility of welcoming you to the Asia Pacific Science Press family. Together, let's forge a future of unparalleled scientific advancement and discovery.

Warm regards

Asia Pacific Science Press

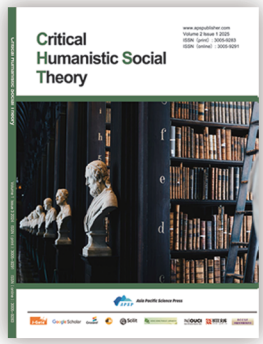
Asia Pacific Science Press

Our Journals

Asia Pacific Economic and Management Review is an international, peer-reviewed and open access journal which focuses on theoretical and applied studies of corporate and financial behavior. Aiming to promote the research in fields of business economics and management, it covers mainly but not limits to the following areas: accounting and financial management, economics, human resource management and organizational behavior, information management, international business, strategy and innovation, management science and operations management, marketing and retailing, finance.



Critical Humanistic Social Theory is an journal that publishes papers specifically using quantitative or qualitative research methods for social science research. The journal encourages scholars to conduct social science theory research from the perspective of social critical theory and emphasizes research concerned with issues or methods that cut across traditional disciplinary lines.



Journal of Educational Theory and Practice is an international, peer-reviewed and open access journal which is to promote the evaluative, integrative, theoretical and methodological research on contemporary education; shape a novel, broader view of issues in contemporary education; enhance the caliber of humanities research through active use of best domestic and foreign practices; and integrate the achievements of various sciences and knowledge areas with unconventional approaches.



Journal of Advances in Engineering and Technology is an international, peer-reviewed and open access journal which publishes original articles, reviews, short communications, case studies and letters in the field of electronic research and application.



Advances in Management and Intelligent Technologies is an international, peer-reviewed, open-access academic journal, hosted by the Fujian Strait Institute of Intelligent Equipment and managed and published by Asia-Pacific Science Press. It focuses on the latest research in the fields of management and intelligent technologies, and aims to advance both theoretical and applied research in management, technological innovation, and intelligent development.



Asia Pacific Journal of Clinical Medical Research is an international, peer-reviewed, open access journal dedicated to advancing clinical medical research across multiple disciplines. The journal serves as a platform for publishing high-quality original research, reviews, and clinical studies that enhance the understanding of medical practices, treatment innovations, and healthcare outcomes, thereby supporting patient care and medical advancements in the Asia Pacific region and beyond.



Asia Pacific Journal of Educational Research is an international, peer-reviewed, open-access academic journal focusing on educational theory and practice. It publishes high-quality research on educational reform, teaching methods, educational equity, and policy studies. The journal addresses practical needs and institutional changes in the education systems of the Asia-Pacific region, advocating a balance between theoretical inquiry and practical experience. It encourages original studies from multicultural, comparative, and interdisciplinary perspectives, aiming to support educational innovation and policy development across the region.



Asia Pacific Economic and Social Development is an international, peer-reviewed, open-access academic journal openly distributed to the global academic community. The journal is committed to publishing original research with theoretical depth and practical value in the fields of economic and social development. It focuses on issues such as economic behavior, social structure transformation, policy innovation, and regional coordinated development in the Asia-Pacific region. The journal encourages interdisciplinary perspectives and promotes the integration of economics, sociology, management, and related disciplines.

