

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 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

Table of Contents

- 1 Research on the Impact of Data Asset Recognition on Corporate Financial Reporting Quality in the Telecommunications Industry—A Case Study of China Mobile**
Zhiwu Wang, Anqi He, Xiang Liu, Yao Tong, Hanyu Wang, Qian Liu, Qilu Liu
- 8 Analysis of the Impact of Artificial Intelligence on Middle-Aged Workers' Employment Willingness: Based on the Context of Delayed Retirement**
Ming Fu
- 16 Axial Compression Effect on Carbonation Resistance of Fly Ash Concrete: An Experimental and Modeling Study**
Hu Li, Zhuang Tian, Fulei Liu, Pengcheng Li
- 23 Research on the Path of Carbon Market Risk Governance Empowered by Accounting Tools**
Xinyao Zhang
- 36 The Impact of the Digital Economy on Municipal Air Quality: Insights from Spatial Analysis**
Xinyan Yang, Lijun Yuan
- 52 An Integrated Financial–Operational Framework for Capital Equipment Decisions: Advancing Manufacturing Investment Theory and Practice**
Hamza Saad
- 65 Why Do Graduates Turn Away from Private Firms? Examining the Roles of Perceived Job Security, Perceived Management Compliance, Perceived Humanistic Care, and Organisational Attractiveness in the Chinese Context**
Weixiang Gan, Mengfei Xiao, Naiqian Zhang, Yongli Zhang, Tara Ahmed Mohammed
- 92 Research on the Evaluation of ESG Practice Effectiveness of Shanxi J Company**
Shijie Sun
- 104 A Bilevel GA–PSO–MILP Framework for Capacity Planning and Day-Ahead Scheduling of an Integrated Electricity–Heat–Cooling Energy System**
Mengzhe Yu, Mingshen Xu, Bomian Cheng

- 119 **Research and Development Trends of Expressway Maintenance Management Technology**
Qingsong Zhang
- 128 **Research on the Construction of Digital Standard Contracts for Artificial Intelligence Data Training**
Qiuhui Ren, Xiaohua Fu
- 135 **Green Credit Policy and the Sustainability of Green Innovation: Symbolic or Substantive Responses from High-Polluting Firms**
Dicheng Wang, Qian Zhang, Ziyue Jin, Han Li
- 152 **Research on the Impact of Long Term Care Insurance on Family Medical Expenditure: Empirical Evidence Based on CHARLS Data from 2011 to 2020**
Churui Yang, Rongxin Li, Xunwen Zheng
- 167 **The Spillover Effect of Financial Agglomeration on Urbanization: An Empirical Study in Southwestern China**
Xinyu Liu
- 178 **Value-Building Rather Than Risk-Reduction? Effects of TikTok Beauty Recommendation Videos on College Students' Purchase Decisions**
Qiuyan Li, Fengyu Zhao
- 192 **Hybrid Course Design for Applied Management Learning in Higher Vocational Colleges: Digital Capability and Student Engagement as Pedagogical Pathways**
Baoyin Liu, Rozaini Binti Rosli
- 201 **Reconstructing Vocational Teacher Capacity in AI-Enabled Industry-Education Integration: A Policy and Institutional Analysis of Liaoning Province**
Chi Tang, Rozaini Binti Rosli
- 211 **Middle Managers and AI-Assisted Performance Appraisal in Chinese Enterprises: Explainability, AI Literacy, and the Case for Human Oversight**
Peng Zhang, Rozaini Binti Rosli
- 219 **Research on the Challenges and Pathways to Breakthroughs in the Development of Consumer Finance in China**
Zirong Liu, Siyuan Yuan

Research on the Impact of Data Asset Recognition on Corporate Financial Reporting Quality in the Telecommunications Industry—A Case Study of China Mobile

Zhiwu Wang, Anqi He, Xiang Liu, Yao Tong, Hanyu Wang, Qian Liu, Qilu Liu*

School of Business, Jiaxing University, Zhejiang Province, 314001, China

*Corresponding author: Qilu Liu, 00008184@zjxu.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: Against the backdrop of a rapidly expanding digital economy and progressively refined regulatory frameworks governing data resources, data assets have assumed an unprecedented strategic prominence as core corporate strategic resources. This study examines the telecommunications industry, with a particular focus on China Mobile as a representative case. Drawing upon the data resource disclosure practices in its 2024 financial report, we systematically investigate the implications of data asset recognition for financial statement presentation and information disclosure. By constructing a financial reporting quality evaluation framework, this research specifically analyzes how the recognition and measurement of data assets influence key qualitative characteristics of financial information, including reliability, relevance, and materiality. Our findings not only contribute to enhanced standardization and transparency in data asset management within the telecommunications sector but also provide a theoretical foundation for firms seeking to optimize financial data governance, unlock data asset value, and strengthen their competitive positioning in the digital economy.

Keywords: Data Assets; China Mobile; Financial Reporting Quality; Accounting Recognition; Information Disclosure; Balance Sheet; Intangible Assets; Digital Economy

Published: Mar 10, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1090>

1. Introduction

The accelerated development of the digital economy has established data as both a fundamental strategic resource at the governmental level and a core factor of production driving value creation at the enterprise level. To regulate the development of markets for data factors, Chinese authorities have successively promulgated a series of policy documents. In August 2023, the Ministry of Finance issued the “Interim Provisions on Accounting for Enterprise Data Resources” (hereafter “Interim Provisions”), stipulating that from January 1, 2024, enterprises may recognize eligible data resources as assets, classifying them as “inventory,” “intangible assets,” or “development expenditure” within financial statements. This regulatory development marks the formal integration of data as the “fifth factor of production” into corporate accounting systems, representing a milestone in advancing the process of data capitalization^[1].

As a quintessential data-intensive sector, the telecommunications industry’s data resources—their scale, quality, and value conversion capacity—directly influence firms’ core competitiveness and sustainable development potential. Within this context, issues concerning the accounting recognition, measurement, and disclosure of data assets not only affect the

transparency of corporate financial information but also exert a profound impact on financial reporting quality. Utilizing China Mobile as a case study, this paper systematically examines the mechanisms through which data asset recognition influences financial reporting quality in telecommunications enterprises, with particular attention to its manifestations across reliability, relevance, and materiality dimensions. Furthermore, we elucidate the pathways through which data assets embody value within financial reporting. Our results offer valuable insights for optimizing data asset management practices in the telecommunications industry, provide theoretical support for relevant policy formulation, and contribute to the industry's pursuit of high-quality development.

2. Literature Review

Existing research on data assets primarily revolves around two core issues: conceptual definition and accounting recognition/measurement, forming a relatively systematic theoretical framework.

Regarding conceptual definition, academic understanding of data assets has evolved through a gradual deepening process. Foundational research in information science (Ackoff, 1989) proposed the “Data-Information-Knowledge-Wisdom” (DIKW) hierarchy, defining data as raw records of entities, events, and their interrelationships. With digital economy advancement, scholars have expanded the conception of data assets across multiple dimensions. Innovative research by Zhu Yangyong et al. (2018) transcended traditional classification boundaries, constructing a unified theoretical framework encompassing digital assets, data assets, and information assets based on common characteristics of physical attributes, existential attributes, and information attributes. This perspective finds resonance at the policy level. The Big Data Technology Standards Promotion Committee (CCSA TC601, 2023) explicitly defines data assets as “resources legally controlled by organizations, electronically recorded, measurable, and capable of creating value” across three dimensions: legal attributes, technical characteristics, and economic value. Notably, the “resources-products-assets” three-stage transformation model proposed by Yao Jianguo et al. (2017) reveals the evolutionary pathway of data from raw resources to strategic assets, providing important theoretical foundations for understanding data asset formation mechanisms.

In the realm of recognition and measurement, establishing ownership represents a critical prerequisite for asset recognition, necessitating clarity regarding data attribution and usage rights. The “Opinions on Building a Data Foundation System to Better Leverage the Role of Data Elements” (“Data Twenty Articles”) issued by the Central Committee of the Communist Party of China and the State Council in December 2022 introduced a tripartite rights separation model—data resource holding rights, data processing and use rights, and data product operation rights—de-emphasizing ownership while stressing use rights to facilitate data circulation^[2]. However, practical implementation of this framework continues to face challenges such as ambiguous delineation of rights and prominent compliance risks (Xiong Qiaoqin and Tang Ke, 2021; Chen Gang et al., 2024)^[3]. Research on measurement methodologies exhibits diversification: traditional approaches, including the cost, income, and market methods, each possess distinct advantages yet suffer from significant limitations^[4]. The cost method offers operational simplicity but ignores timeliness and fails to reflect future benefits; the income method aligns with the nature of assets' future economic benefits but relies excessively on subjective forecasts; the market method depends on fair value in trading markets, but data non-standardization results in a scarcity of comparable cases (MIIT, 2023)^[5]. With technological progress, measurement techniques leveraging big data, machine learning, and deep learning are gaining traction. Innovative valuation technologies, exemplified by Shangshu Net's “Golden True Valuation,” integrate deep learning algorithms (e.g., the DeepSeek model) to enable dynamic assessment of data's immediate, strategic, and risk values, offering technical solutions to valuation dilemmas.

In summary, academic consensus has largely coalesced around the connotation and essential characteristics of data assets. They are widely regarded as non-monetary strategic resources, electronically recorded, legally controlled by organizations, transformed through a three-stage “resource-product-asset” process, and possessing measurability and value-creating capacity. Nevertheless, significant divergence persists in critical areas such as measurement standards and recognition methods, highlighted by three major practical challenges: first, the design of rights confirmation systems requires urgent refinement, as the existing “tripartite rights separation” framework lacks detailed operational specifications; second, a valuation standard system remains underdeveloped, with traditional measurement methods struggling to accommodate

data assets' unique attributes^[6]; third, a balancing mechanism between security and development is absent, creating tension between data circulation efficiency and compliance requirements. These challenges constitute important directions for future research.

3. Data Asset Profile of China Mobile

As a leading enterprise in China's telecommunications sector, China Mobile's business scope encompasses mobile communication services, digital services, value-added services, and other areas. Leveraging its vast user base and extensive communication network infrastructure, the company has accumulated substantial data resources. These primarily comprise core data types such as user consumption behavior data, location trajectory data, and network usage data, providing a critical foundation for supporting digital economic development and fostering social service innovation.

In accordance with the Ministry of Finance's Interim Provisions, China Mobile incorporates data resources meeting the asset definition into its financial statements. Initial measurement primarily employs the cost method, specifically encompassing cost aggregation from data collection, cleansing and processing, storage and maintenance, and related platform construction. For internal development expenditures satisfying capitalization criteria, the company records them under the "Development Expenditure - Data Resources" account, transferring their costs to the "Intangible Assets - Data Resources" account upon formation of identifiable data assets. For subsequent measurement, the cost model is applied: data resources recognized as intangible assets are amortized over their expected benefit periods using a systematic and rational method, with impairment testing conducted regularly per accounting standard requirements.

By compiling statements from 2021-2024, China Mobile's relevant data resource presentations are as follows:

Table 1. Overview of China Mobile's Data Resource Incorporation (Unit: CNY Million)

age	2021	2022	2023	2024
floating assets				
total of current assets	595371	456371	498104	568559
non-liquid asset				
intangible assets	44319	46509	47597	50804
Among them: data resources	0	0	0	560
development expenditure	919	1334	2279	2157
Among them: data resources	0	0	0	56
Other non-current assets	1165418	1396024	1409377	1451307
Total non-current assets	1210656	1443867	1459253	1504268
Total assets	1806027	1900238	1957357	2072827

Data source: based on China Mobile's 2021-2024 annual report.

During the process of data asset incorporation, China Mobile has progressively established an accounting framework adapted to the characteristics of new production factors. This not only reflects the transformation of data assets from resources to assets but also provides a reference sample for the industry in exploring the market-oriented allocation of data elements.

As evidenced in Table 1, the absence of uniform accounting standards for data resources from 2021 to 2023 meant that data resources were not separately presented in China Mobile's financial reports. This situation changed significantly with the implementation of the Interim Provisions effective January 1, 2024. The 2024 annual financial report indicates that China Mobile has conducted standardized accounting recognition and presentation of data resources in accordance with the new regulations: the amount recognized as intangible assets reached CNY 560 million, while data resources included in development expenditure amounted to CNY 56 million, yielding a total recognized value of CNY 616 million. This accounting practice signifies the formal representation of data elements as a new asset class within financial statements.

4. Impact of China Mobile's Data Asset Recognition on Financial Reporting Quality

4.1 Impact on Statement Presentation

(1) Optimization of Balance Sheet Structure

According to disclosures in China Mobile's 2024 annual report, incorporated data assets totaled CNY 616 million, with CNY 560 million classified under "Intangible Assets" as "data resources" and CNY 56 million included in "Development Expenditure" as "data resources." Following legal rights confirmation and value assessment, these data assets were formally incorporated into the balance sheet, significantly increasing the book value of intangible assets. This not only optimizes the asset structure of the balance sheet to better reflect the firm's core resource composition in the digital economy era but also elevates total asset scale. Post-incorporation, China Mobile's total assets increased, while total liabilities remained relatively stable or increased less proportionally than assets. The debt-to-asset ratio declined from 33% on January 1, 2024, to 32% on June 30, 2024. This change indicates that data asset recognition improves the firm's capital structure, reduces financial leverage, and a lower debt-to-asset ratio signifies enhanced long-term solvency and financial stability, potentially boosting creditor confidence and reducing future financing costs^[9].

(2) Periodic Smoothing of the Income Statement and Improved Cost Matching

Prior to data asset incorporation, expenditures related to data resource acquisition and development—such as data collection, cleansing, platform construction, and algorithm R&D—were typically expensed fully in the income statement during the current period. This accounting treatment not only understated current-period profits but also failed to accurately reflect the long-term economic value of data resource investments. Following incorporation, relevant costs meeting capitalization criteria are no longer fully expensed immediately but are recognized as assets, classified as "development expenditure" or "intangible assets" based on their nature, and systematically amortized over their expected benefit periods. This new accounting approach more reasonably matches costs with anticipated future revenues, avoiding the concentrated burden of upfront expenses. Particularly during the initial phase of large-scale data asset capitalization, this treatment demonstrably optimizes the income statement.

The incorporation of data assets not only enhances financial statement presentation but, more importantly, promotes operational efficiency in core business activities. By formally recognizing data resources as assets, enterprises can develop and utilize their latent value more systematically. Specific manifestations include: customer profiling analysis based on data assets significantly improves marketing precision; network optimization applications reduce operating costs; risk control models enhance business security; and new product development accelerates innovation. These applications directly drive growth in main business revenue and improvements in profitability. China Mobile's H1 2024 financial data substantiates this point; the synchronous improvement in its operating revenue margin and net sales profit margin is largely attributable to operational efficiency gains stemming from data asset applications.

From a long-term development perspective, data asset incorporation enables the income statement to present a more comprehensive and truthful reflection of the complete value creation process of data resources—from initial input to final output—thereby highlighting the core value of data as a new factor of production in financial performance. This accounting reform not only optimizes current financial metrics but also lays a solid foundation for the firm's future digital transformation.

(3) Significant Expansion in Information Disclosure Depth and Breadth

China Mobile is required to elaborate on the key accounting policies and significant judgments underpinning its recognition and measurement, including the specific application of data asset recognition conditions, cost components of initial measurement, the chosen subsequent measurement model, the amortization method employed and the basis for estimating its useful life, as well as criteria and testing methods for identifying impairment indicators. Links involving significant accounting estimates—such as determining whether development phase expenditures meet capitalization criteria, the basis for establishing the economic useful lives of data assets, and key assumptions underlying future cash flow projections in impairment testing—require adequate disclosure.

This multi-tiered, systematic information disclosure mechanism effectively mitigates information asymmetry inherent to data assets as a new factor of production. The intangible nature, value volatility, and technical complexity of data assets

often make it challenging for external stakeholders to accurately assess their true value. By establishing a robust information disclosure framework, China Mobile enables investors, creditors, and other stakeholders to comprehend—through financial statement data—the scale characteristics, quality levels, management efficiency, and actual contribution to value creation of the firm’s core data resources, thereby enhancing the credibility of financial reports and their decision-usefulness.

4.2 Impact on Financial Reporting Quality

4.2.1 Positive Impacts

(1) Enhancement of Completeness

In the digital economy era, data pertaining to user behavior, location, and online activities have become core strategic resources underpinning China Mobile’s 5G services, digital services, and computing power network, yet remained off-balance-sheet for an extended period due to the absence of accounting standards. With the standardization of data asset accounting, this situation has improved markedly. According to the 2024 annual report, China Mobile recognized data assets valued at CNY 560 million as intangible assets (primarily comprising user identity data and communication service data), while incorporating CNY 56 million of data resources into the development expenditure account (mainly involving network operation data and platform interaction data). The on-balance-sheet presentation of these data assets expands the coverage of intangible assets in traditional financial reporting, offering a more comprehensive reflection of the firm’s new resource value in the digital age. For instance, user location data supports base station optimization and smart transportation services; its incorporation allows investors to more clearly identify the asset base underlying China Mobile’s “connectivity + computing power + capability” integration strategy, significantly improving financial reporting completeness and better aligning with the “substance over form” principle (Zhang Anling, 2025)^[7]. At a deeper level, the on-balance-sheet presentation of data assets provides a crucial practical pathway for addressing the insufficient relevance of accounting information in the digital economy era.

(2) Strengthening of Relevance

Data asset incorporation has significantly enhanced the information relevance of China Mobile’s financial reporting, enabling it to better meet the decision-making needs of various stakeholders (Huang Can et al., 2024)^[7]. From an external stakeholder perspective, China Mobile’s data assets are directly linked to its core competitive advantages: user behavior data supports precision marketing to increase average revenue per user, while network operation data aids in optimizing base station resource allocation to reduce operating costs. Once the value of these data assets is manifested through the financial reporting system, investors can more accurately assess the firm’s value creation capacity within the digital economic environment. Taking the CNY 560 million in intangible assets recognized in 2024 as an example, user data assets related to the “Smart Family” business accounted for 30% of the total—a key piece of information providing investors with an important basis for forecasting this business segment’s future growth potential.

From an internal management standpoint, the accounting information generated post-incorporation holds significant decision-support value. Specifically, the capitalization of data asset development costs enables the amortized cost over the asset’s useful life to form a matching relationship with the value-added service revenue generated by corresponding business activities. By analyzing the dynamic matching between amortized costs and multi-period business revenues, management can more scientifically optimize the allocation direction of data resources, thereby establishing a virtuous feedback mechanism of “data input-business output.” This accounting approach not only enhances the scientific rigor of management decisions but also strengthens the guiding role of financial information in corporate strategy implementation.

(3) Reflection of Materiality

Data asset incorporation underscores the materiality of data resources within China Mobile’s value creation system. In traditional financial reporting, data-related resources were often broadly categorized under “intangible assets” or expensed, failing to reflect their strategic value. With the formal establishment of data as the “fifth factor of production,” China Mobile’s data assets have become deeply embedded within its “connectivity + computing power + capability” strategic framework: a base station layout optimization system based on user location data reduces annual network operating costs by approximately CNY 200 million; value-added business revenue generated by content recommendation services relying on online behavior

data analysis exceeded CNY 5 billion in 2024. By recognizing such strategic data resources as CNY 560 million in intangible assets and CNY 56 million in development expenditure, China Mobile clearly conveys a critical signal to the capital market: “data resources constitute core corporate competitiveness.” This not only prevents investors from undervaluing its digital business potential but also enables financial reports to more accurately reflect the composition of the firm’s core economic resources, highlighting the materiality of data assets for long-term development.

4.2.2 Negative Impacts

(1) Impact on Reliability

The inherent characteristics of data assets render their valuation challenging, thereby posing threats to financial reporting reliability. Measurement difficulties are primarily manifested in several aspects: First, the initial cost aggregation for data assets is complex, particularly for self-developed assets, where the capitalization judgment for R&D expenditures involves a degree of subjectivity. Second, subsequent measurement of data assets also faces numerous issues; for instance, fair value measurement relies on complex valuation models and parameter assumptions, while the lack of market transaction data impairs the reliability of valuation outcomes (Yao Shuo, 2024)^[10]. Furthermore, the value of data assets is susceptible to factors like technological iteration and shifts in market demand, exhibiting high dynamism and uncertainty. These factors may lead to discrepancies between the book value and actual value of data assets, potentially impairing financial reporting reliability.

(2) Impact on Comparability

Data asset incorporation presents a dual challenge to financial reporting comparability. Horizontally, significant disparities in accounting treatment exist within the industry. For example, China Unicom (2024) included 80% of its data assets in the development expenditure account, whereas China Mobile recognized 91% of similar assets as intangible assets. Regarding measurement methods, China Telecom’s adoption of the income method contrasts sharply with China Mobile’s cost method, resulting in a lack of comparable basis for asset values among peers. Vertically, the data asset information generated by China Mobile’s initial implementation of the Interim Provisions in 2024 cannot be linked to historical periods and may become disjointed from subsequent periods due to future standard revisions or measurement method improvements (e.g., introduction of fair value). More notably, its disclosure framework fails to distinguish between the value of raw data and value-added data products. This inadequate disclosure further undermines the comparability of financial information across firms.

4.3 Impact on Corporate Valuation

As a pivotal institutional innovation in the digital economy era, data asset incorporation is systematically reshaping the corporate valuation framework. Its internal financial optimization effects operate through a dual pathway: On one hand, data resources meeting recognition criteria are incorporated into the balance sheet, directly expanding total assets. For instance, the disclosed scale for 92 A-share listed firms that incorporated data assets in 2024 reached CNY 2.495 billion, a thirty-fold increase from the beginning of the year. On the other hand, capitalization replaces expensing, significantly reducing current-period expenses in the income statement; for example, manufacturing firms optimize profit structures through data asset amortization. Research by Xu Pan and Li Jieyi (2024)^[11] demonstrates that this dual-statement linkage effect not only improves the debt-to-asset ratio and increases the return on equity (ROE) but also effectively mitigates the “hollowing out” problem of asset-light firms’ balance sheets, steering valuation toward greater accuracy and rationality.

Regarding external valuation, data asset incorporation enhances external valuation accuracy by optimizing financial representation and strengthening market expectations^[12]. At the capital structure level, recognizing data resources as assets directly lowers the debt-to-asset ratio, boosting creditor confidence and broadening financing channels; optimization of solvency indicators further strengthens credit ratings and reduces debt costs. In terms of profitability, ROE improves through dual pathways—expansion of net asset scale and optimization of profit structure. Operationally, data asset-driven refined operations reduce inventory turnover days and accounts receivable cycles, with improvements in total asset turnover signaling enhanced operational efficiency to the market. At the development capacity level, the scale and value-added potential of data assets are incorporated into growth expectations. In summary, the incorporation system, through increased transparency of financial indicators and quantifiability of future earnings, enables capital markets to price data elements more accurately.

5. Conclusion

Focusing on the telecommunications industry and employing China Mobile as a Typical case, this paper systematically reveals the dual impact mechanism of data asset recognition on financial reporting quality. The study finds that the capitalization of data assets significantly optimizes the balance sheet structure and achieves periodic smoothening of the income statement through a deferred amortization mechanism. The impact of data asset incorporation on financial reporting quality is twofold: it markedly enhances completeness, relevance, and materiality, yet, due to disparities in accounting policies, may also negatively affect reliability and comparability. This research not only deepens the theoretical understanding of data asset recognition but also provides practical guidance for the telecommunications industry to optimize data asset management and improve the quality of financial information disclosure^[13].

Funding

Thanks to the Regular Project Fund of Zhejiang Provincial Philosophy and Social Sciences Planning (No.25NDJC081YB), the Statistical Research Project of Zhejiang Province (No.25TJYX29) and the Special Project Fund of Jiaxing University's Humanities and Social Sciences "Young and Outstanding Talent" Cultivation Program (No.007JKY155AW) for their support of this research.

Conflict of Interests

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

Reference

- [1] Zhu Y. Y., & Ye Y. Z. (2018). From the perspective of the attributes of data, data assets. *Big Data*, 4(06), 65-76.
- [2] Xiong, Q. Q., & Tang, K. (2021). Research Progress on the Definition of Rights, Transaction, and Pricing of Data Elements. *Economic Perspectives*, (2), 143-158.
- [3] Chen, G., Yan, B. B., & Tang, K. (2024). Data Elements and Assets: Theoretical Analysis and Practical Exploration. *International Economic Review*, (5), 153-176+8.
- [4] Kim, O. K., Park, J., Park, C. W., et al. (2021). Data asset valuation model review. *The Journal of Big Data*, 6(1), 153–160. <https://doi.org/10.1186/s40537-021-00435-9>
- [5] Ministry of Industry and Information Technology (MIIT). (2023). White Paper on Data Asset Valuation in the Telecommunications Industry (2023 Edition).
- [6] Mu, M. D., Li, J. P., & Hao, J. (2025). The Incorporation of Data Assets into Financial Statements: Framework, Pathway and Practical Exploration. *Procedia Computer Science*, 266, 1109–1115. <https://doi.org/10.1016/j.procs.2025.08.137>
- [7] Qian, K., Pan, C. X., & Liang, X. Y. (2025). Does data asset disclosure affect bank lending acquisition? *International Review of Financial Analysis*, 97, 103834. <https://doi.org/10.1016/j.irfa.2024.103834>
- [8] Zhang, A. L. (2025). The Impact of Data Asset Recognition on the Financial Reporting Quality of Rail Transit Enterprises. *Business Watch*, 11(12), 36-39+43.
- [9] Huang, C., Shi, Z. Q., & Jiang, Q. S. (2024). The Impact of Data Asset Recognition on Enterprises: Evidence from the Stock Market. *Management Science*, 37(6), 43-61.
- [10] Yao, S. (2024). Research on the Impact of Data Asset Recognition in the Digital Economy on Corporate Financial Reporting Quality. *Chinese Agricultural Accounting*, 34(22), 73-75.
- [11] Xu, P., & Li, J. Y. (2024). Pathways for Enterprise Data Asset Recognition: Framework and Practice. *Accounting Monthly*, 45(7), 58-62.
- [12] Li, Y., Luo, C. K., Dong, L. M., & Gui, M. Z. (2022). Data Asset Disclosure and Nonprofessional Investor Judgment: Evidence from Questionnaire Experiments. *Mobile Information Systems*, 2022, 8116063. <https://doi.org/10.1155/2022/8116063>
- [13] Zhang, J. R., Wei, Y. L., & Song, X. Y. (2020). Research on Accounting Treatment and Information Presentation of Enterprise Data Assets. *Accounting and Economic Research*, 34(3), 1-16.

Analysis of the Impact of Artificial Intelligence on Middle-Aged Workers' Employment Willingness: Based on the Context of Delayed Retirement

Ming Fu*

Guangzhou Institute of Science and Technology, Guangzhou, 510540, China

*Corresponding author: Ming Fu, 15796734137@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 rapid development of AI and China's delayed retirement policy have significantly challenged middle-aged workers' employment willingness. This study utilized a cross-sectional survey of 889 pre-retirement individuals in Beijing, Guangzhou, and Lanzhou, using multivariate regression analysis to examine key influencing factors. Results indicate that employment willingness is significantly higher among males and highly educated individuals, while widespread AI adoption in eastern and northern regions increases pressure on low-educated groups. Notably, household economic pressure correlates negatively with work intentions. The study concludes that AI's impact varies across demographics, necessitating targeted vocational training and social support to help middle-aged workers adapt to the modern job market.

Keywords: Artificial Intelligence; Delayed Retirement; Middle-Aged Workers; Employment Intention; Regional Differences; Multivariate Regression; Vocational Training

Published: Mar 10, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1103>

1. Introduction

China's implementation of the delayed retirement policy has been seen by many scholars as unexpected, especially under the dual pressures of rapid AI development and economic downturn. The employment willingness of middle-aged workers nearing retirement has garnered widespread attention. The potential threat of AI to jobs, along with the challenges posed by technological advancements for middle-aged and older workers, is becoming a focal point for policymakers and researchers. As noted by the Oxford Institute of Population Ageing, the introduction of AI requires workers, especially older employees, to continuously relearn new skills in order to adapt to the digital transformation of the work environment (Ferdous, 2023). Meanwhile, the delayed retirement policy aims to ease labor market pressures associated with population aging and to encourage older individuals to remain economically active (Haan & Tolan, 2019).

However, despite the new employment opportunities potentially created by delayed retirement policies and the application of AI technology, the reemployment willingness of middle-aged workers is not always positive. On the one hand, many middle-aged individuals face considerable pressure to update their skills for reemployment, particularly in adapting to AI-driven work environments, where they often encounter greater difficulties than younger workers. On the other hand, the combined pressures of technological change and later-life labor participation may intensify insecurity among middle-aged and older workers, rather than automatically translating into stronger willingness to remain in or re-enter the labor market.

Artificial intelligence (AI) technology is advancing at an astonishing pace and is being widely applied across various sectors, particularly in the employment field. For middle-aged workers, the rapid development of AI brings challenges in skill adaptation and reshapes the labor market. In this context, China has introduced a new delayed retirement policy to address labor shortages caused by an aging population. In recent years, the application of AI in the job market has significantly altered the nature of work, including the automation of repetitive tasks, changes in skill requirements, and the reconfiguration of workers' job experiences and opportunities (Bankins et al., 2024).

The aim of this study is to analyze the impact of AI on the employment intentions of middle-aged workers in the context of delayed retirement. First, this study will explore how AI affects employment choices by increasing skill requirements and reshaping the labor market. Second, it will examine how the combination of delayed retirement policies and AI application influences the work motivation of middle-aged individuals. As recent research suggests, older and aging workers in rapidly evolving digital environments often face substantial learning and adaptation pressures, making lifelong learning and technological support increasingly important for sustaining employability (Ranasinghe et al., 2024). Finally, this study seeks to provide policy recommendations to help middle-aged workers better adapt to AI-driven job markets.

2. Literature Review

2.1 The Impact of AI Technology on Middle-Aged Workers

Artificial intelligence (AI) technology is rapidly transforming the global job market, and its influence has garnered considerable attention from scholars. Recent research suggests that AI is reshaping work by altering task structures, changing skill requirements, and reconfiguring workers' experiences within organizations, while also creating demand for more advanced and adaptive capabilities (Úbeda-García et al., 2025). However, for middle-aged workers, adapting to these changes is often challenging. Research on aging and older workers shows that technological transformation frequently generates substantial learning and adaptation pressures, especially where new digital competencies must be acquired continuously in order to remain employable. In addition, evidence from workers aged 40 and above indicates that the availability of digital technology training plays an important role in supporting their intention to remain in work, suggesting that insufficient or poorly tailored training can deepen their disadvantage in technology-intensive labor markets (Xie et al., 2023). The skill gap and limits in technological adaptation may therefore restrict career development opportunities and weaken willingness to seek reemployment or career transitions among middle-aged workers.

The decline in employment willingness is not only reflected in difficulties with technical learning but also in the psychological resistance middle-aged workers face when adapting to new work environments. With the introduction of AI technology, workplaces have become increasingly digitalized and complex, and this transformation can generate substantial pressure for workers who must continuously adjust to new tools, routines, and skill demands. For middle-aged workers accustomed to more traditional work models, such changes may create stronger adaptation burdens and uncertainty. Research on aging workers shows that continuous technological change often places older employees under considerable learning pressure, especially when upskilling becomes necessary to remain employable (Chang et al., 2023). At the same time, studies of digital transformation indicate that employee resistance to change is closely linked to stress and reduced well-being, which can further weaken motivation to engage in retraining or workplace transition (Valtonen et al., 2024). In addition, evidence on workers aged 40 and above suggests that when digital training is insufficient or not tailored to their needs, their confidence in continued work participation may decline (Xie et al., 2023). As a result, emotional strain and adaptation difficulties may further diminish middle-aged workers' motivation to participate in reemployment or skill-upgrading programs.

2.2 The Impact of Delayed Retirement Policies on Worker Reemployment

The implementation of delayed retirement policies is intended to respond to the economic and demographic pressures associated with population aging, including labor supply pressures and the sustainability of long-term development. The rationale behind this policy is to extend working lives and maintain labor market participation in aging societies (Zhao et al., 2024). However, the implementation of delayed retirement policies has also encountered significant challenges. Recent research on China indicates that delayed retirement has complex effects on labor force participation and broader socioeconomic outcomes, suggesting that policy extension alone does not automatically translate into stronger employment

engagement among older and middle-aged workers (Dai et al., 2025).

Research also shows that many middle-aged workers have reservations about delayed retirement policies. These concerns are related not only to skill renewal, but also to whether adequate digital training and work-related support are available at this career stage. Evidence from workers aged 40 and above suggests that the availability of digital technology training can significantly strengthen older workers' intention to remain in work, especially when such training enhances their sense of competence and usefulness in increasingly technology-intensive workplaces. This suggests that when training support is insufficient or poorly aligned with the needs of middle-aged workers, delayed retirement policies may have limited success in increasing their motivation to remain employed or re-enter the labor market.

2.3 The Limitations of the Policy Framework and the Unique Needs of Middle-Aged Workers

The current policy framework remains insufficient in supporting middle-aged workers in adapting to AI-driven employment changes. Recent research suggests that digitalized work environments can heighten older workers' vulnerability to stress, exclusion, and adaptation difficulties when support systems are weak, and that targeted digital literacy initiatives together with sustained managerial engagement are crucial for ensuring that technological change enhances rather than undermines their well-being and productivity.

To better support middle-aged workers, policymakers must consider their unique needs. This includes providing more flexible learning and work arrangements, as well as strengthening organizational and social support. For instance, psychological counseling and career guidance services can help middle-aged workers manage the stress and anxiety associated with career transitions. In addition, evidence from workers aged 40 and above shows that digital technology training can significantly strengthen intentions to remain in work, especially when such training is perceived as useful and aligned with workers' developmental needs.

2.4 The Future Direction of Integrating Policy and Technology

As AI technology continues to advance, policymakers must consider how to integrate technological change into more adaptive labor policies. Recent research suggests that the contemporary labor market is being reshaped not only by AI and related technological change, but also by shifts in where, how, and under what conditions work is performed, which in turn requires more age-sensitive and flexible labor and human resource practices across different career stages (Truxillo et al., 2026). In this context, policies should place greater emphasis on inclusive career development, targeted upskilling, and flexible work arrangements that can help middle-aged workers navigate career transitions more smoothly. In addition, governments may consider using economic incentives and supportive welfare arrangements to reduce the burden of retraining and encourage employers to provide more transition-friendly employment opportunities (Wang et al., 2024).

A review of the existing literature reveals that the rapid development of AI technology, alongside the implementation of delayed retirement policies, is profoundly affecting middle-aged workers' employment prospects and career development. While AI may create new opportunities, it can also raise barriers to labor market adaptation for workers who face greater challenges in responding to technological change. The current policy framework therefore needs to become more flexible and better aligned with technology-driven changes in work organization and skill requirements. By strengthening targeted training, improving age-inclusive work design, and offering supportive transition policies, policymakers can help middle-aged workers better integrate into the evolving job market.

3. Methods

This research aims to explore the impact of artificial intelligence (AI) on the employment intentions of pre-retirement individuals. The primary objective is to assess changes in employment attitudes across different age groups as they face advancements in AI technology. The sample includes three major cities in China: Beijing, Guangzhou, and Lanzhou, targeting individuals aged 50 and above. The survey covered key areas such as personal information, employment status, attitudes toward delayed or early retirement, and AI technology usage. Trained staff at mobile monitoring centers guided participants in completing the questionnaire, explaining specific requirements like self-reported age. In total, 1,000 pre-retirement individuals were surveyed, with 889 valid responses. Female respondents were near 50 years old, while males were near 60. Female retirees who experienced discrepancies due to flexible retirement mechanisms were excluded to ensure sample

validity and representativeness.

3.1 Study Design

This study uses a cross-sectional survey design to analyze the impact of AI technology on the employment intentions of pre-retirement individuals. As AI technology reshapes the labor market, particularly for those nearing retirement, this research employs quantitative methods to systematically assess changes in employment attitudes across different age groups and genders. The study is based on cross-sectional survey data and uses multivariate regression analysis to examine the relationships between AI-related factors, demographic characteristics, and employment intentions. Key research areas include personal information, employment status, attitudes toward delayed or early retirement, and knowledge and usage of AI technology.

3.2 Sample

Participants were recruited through site-based field survey sampling to achieve regional variation and practical coverage of pre-retirement populations. The sample was drawn from three representative cities in China's southern, northern, and western regions: Guangzhou, Beijing, and Lanzhou. These locations were selected to capture variation in economic development, labor market conditions, and exposure to AI-related technological change. The target population consisted of individuals aged 50 and above who were approaching retirement and therefore facing potential employment and retirement-related transitions. A total of 1,000 pre-retirement individuals were approached for participation, of whom 889 provided valid responses. Female respondents were generally near the statutory retirement age of 50, while male respondents were generally near the statutory retirement age of 60. To reduce confounding influences related to heterogeneous retirement arrangements, female respondents whose retirement status was substantially affected by flexible retirement policies were excluded from the final analysis. By covering multiple regional contexts and a diverse group of pre-retirement individuals, the sample provides a useful basis for examining employment intentions under conditions of delayed retirement and AI-related labor market change.

3.3 Data Collection

Data were collected through face-to-face questionnaire surveys administered by trained research staff in field-based survey settings. Survey sites were established in community service centers and senior activity venues, where pre-retirement individuals commonly gather, in order to facilitate access to the target population across the three cities. This approach helped improve response quality by allowing participants to complete the questionnaire in supervised settings while retaining independence in their responses.

Before administration, research staff provided a standardized explanation of the study purpose and questionnaire structure, and clarified the meaning of specific questions when necessary. However, all questionnaires were completed independently by participants in order to minimize interviewer interference and avoid proxy responses. The survey covered demographic characteristics (e.g., age, gender, education level, and region), employment history, retirement plans, attitudes toward delayed and early retirement, and respondents' awareness and use of AI technologies. This procedure was designed to enhance data accuracy, consistency, and reliability across survey locations.

3.4 Measures and Variables

The primary measurement instrument was a structured questionnaire developed for this study with reference to established measures used in related research on employment intention, retirement attitudes, and technology use. To improve measurement quality, the questionnaire was pilot-tested prior to formal data collection, and internal consistency was assessed using reliability analysis (e.g., Cronbach's alpha where applicable). Minor revisions were made following the pilot stage to improve clarity and item wording.

The key independent variables included demographic and contextual characteristics, such as age, gender, education level, region, occupation type, and family economic pressure, as well as AI-related factors, including AI awareness and frequency of AI use. The questionnaire also included items on retirement attitudes and retirement planning in order to examine how respondents positioned themselves with regard to delayed or early retirement. The dependent variable was employment intention, operationalized through respondents' reported willingness to remain in work, seek re-employment, engage in part-time work after retirement, or continue full-time employment where feasible. These measures allowed the study to assess how

individual characteristics, regional context, and AI-related factors were associated with employment intentions among pre-retirement individuals.

3.5 Data Analysis

Data analysis was conducted using Stata. The analytical procedures included descriptive statistics, correlation analysis, and multivariate regression models to examine how demographic characteristics, AI awareness, AI usage frequency, region, occupation type, and family economic pressure were associated with employment intentions. The analysis focused on identifying the relationships between individual characteristics and employment intentions under conditions of delayed retirement and AI-related labor market change.

In the regression models, gender, age, education level, region, occupation type, family economic pressure, AI awareness, and AI usage frequency were included as key explanatory variables. Additional control variables, such as health status, were incorporated to improve the robustness of the results. Interaction terms were further introduced to examine whether the effects of gender and education on employment intentions varied across regions. This analytical strategy allowed the study to assess both the direct associations of key variables and the moderating role of regional context.

3.6 Ethical Considerations

The study design and data collection procedures adhered to standard ethical principles for social science research. Before completing the questionnaire, all participants were informed of the purpose of the study, the voluntary nature of participation, the confidentiality of their responses, and their right to withdraw at any stage without consequence. Written informed consent was obtained from all participants prior to data collection.

To protect participants' privacy, all questionnaire data were anonymized prior to analysis, and no personally identifiable information was included in the final dataset. The data were used exclusively for academic research purposes and were not disclosed to any third party. These procedures were implemented to ensure privacy protection, confidentiality, and ethical compliance throughout the research process.

4. Results

4.1 The Impact of Gender and Education on Employment Intentions

This study first examined the associations of gender and education with employment intentions among pre-retirement individuals. The regression results showed that men reported significantly higher employment intentions than women ($\beta = 0.48$, $p < 0.01$). In addition, individuals with higher levels of education (bachelor's degree or above) exhibited significantly stronger employment intentions than those with lower educational attainment (high school or below) ($\beta = 0.52$, $p < 0.01$). These findings indicate that, in the context of delayed retirement, gender and education are significantly associated with willingness to remain in or return to employment. Specifically, male respondents and more highly educated individuals appeared more likely to express continued labor market participation intentions. The detailed regression results are presented in Table 1.

Table 1. Regression Analysis Results of Gender and Education on Employment Intentions

Variable	Regression Coefficient (β)	Standard Error (SE)	Significance Level (p)
Gender (Male)	0.48	0.11	< 0.01
Education (Higher Education)	0.52	0.10	< 0.01

4.2 Regional Differences Analysis

The regression results further suggest that employment intentions varied across regions. In particular, respondents with lower education levels in the eastern and northern regions showed significantly lower employment intentions, indicating stronger pressure related to retirement and labor market adjustment ($\beta = -0.45$, $p < 0.05$). At the same time, employees in the eastern region reported significantly stronger concerns about job replacement associated with AI-related technological change ($\beta = 0.40$, $p < 0.05$). However, respondents in these regions also demonstrated a stronger willingness to improve their competitiveness through digital skills training ($\beta = 0.58$, $p < 0.01$). Taken together, these findings suggest that regional context plays an important role in shaping how workers perceive and respond to labor market changes associated with AI.

Compared with other regions, individuals in the eastern and northern regions appear more likely to experience both higher perceived pressure and stronger adaptive responses, especially through skill development. Compared with other regions, individuals in the eastern and northern regions appear more likely to experience both higher perceived pressure and stronger adaptive responses, especially through skill development. These findings suggest that regional context moderates how gender and education are related to employment intentions.

4.3 The Role of Occupation Type and Family Economic Pressure

Occupation type was also significantly associated with employment intentions. Respondents in technical occupations showed stronger employment intentions than those in non-technical positions ($\beta = 0.32$, $p < 0.01$). This finding suggests that workers in technical roles may feel better positioned to adapt to ongoing technological change, possibly because they possess stronger digital competencies or perceive lower risks of displacement. By contrast, non-technical workers may face greater uncertainty regarding future employment, which may weaken their willingness to remain in the labor market.

Family economic pressure was another significant factor. The results indicate that greater family economic pressure was negatively associated with employment intentions ($\beta = -0.39$, $p < 0.01$). This finding suggests that financial strain may not always encourage continued employment; instead, under some circumstances, it may intensify stress and reduce individuals' willingness to remain in or re-enter work. The detailed regression results are presented in Table 2.

Table 2. Regression Analysis of Occupation Type and Family Economic Pressure on Employment Intentions

Variable	Regression Coefficient (β)	Standard Error (SE)	Significance Level (p)
Occupation Type (Technical)	0.32	0.09	< 0.01
Family Economic Pressure	-0.39	0.08	< 0.01

4.4 Interaction Effect Analysis

This study further examined whether the effects of gender and education on employment intentions differed across regions. The interaction analysis showed that the combined effect of gender, education, and region was significant (interaction term $\beta = 0.22$, $p < 0.05$), suggesting that the relationship between individual characteristics and employment intentions is shaped by regional context.

More specifically, in the eastern region, men with higher education levels reported significantly stronger employment intentions than other groups, whereas women with lower education levels showed comparatively lower employment intentions. This pattern suggests that the advantages associated with education and gender may be more strongly reflected in regions characterized by faster technological development and more intensive labor market transformation. By contrast, in the western and central regions, the interaction effects of gender and education were less pronounced, indicating that regional economic conditions and the pace of technological adoption may moderate how these factors are related to employment choices. The main coefficients for the regional and interaction models are summarized in Table 3.

Table 3. Summary of Regional and Interaction Regression Results on Employment Intentions

Variable	Regression Coefficient (β)	Standard Error (SE)	Significance Level (p)
Lower education in eastern and northern regions	-0.45	0.18	< 0.05
Concern about AI-related job replacement in the eastern region	0.40	0.17	< 0.05
Willingness to improve competitiveness through digital skills training	0.58	0.14	< 0.01
Gender \times Education \times Region	0.22	0.10	< 0.05

Note. The table reports the main coefficients from the regional and interaction regression models. Positive coefficients indicate a stronger employment intention or stronger related attitudinal tendency, whereas negative coefficients indicate lower employment intention or greater retirement-related pressure.

5. Discussion

The findings of this study offer a significant theoretical advancement by elucidating the mechanisms through which Artificial Intelligence (AI) reconfigures labor market experiences, particularly through the framework of Social Cognitive Theory (SCT). SCT posits that human behavior is a product of “reciprocal determinism” involving personal factors, environmental events, and self-efficacy. In this research, AI is identified as a disruptive environmental catalyst that profoundly erodes the “self-efficacy” of middle-aged workers. When AI is perceived not as an augmentative tool but as an insurmountable existential threat to established skills, it diminishes workers’ perceived capability to navigate the labor market successfully. This psychological impingement is further exacerbated by the delayed retirement policy, which extends the temporal horizon of this technological confrontation, essentially tethering workers to a digitalized infrastructure they feel ill-equipped to master. Crucially, the paradoxical finding—that heightened household economic pressure negatively predicts employment willingness—is theorized here as a localized manifestation of “learned helplessness” within the SCT paradigm. While traditional economic models view financial strain as a primary driver for labor participation, the “dual impingement” of AI-driven displacement and mandatory career extension may induce a psychological state of perceived futility. For middle-aged individuals facing significant financial burdens, the inability to control technological shifts can lead to a collapse of “outcome expectations.” In such instances, the perceived lack of agency results in a maladaptive coping mechanism: rather than intensifying labor effort, individuals may opt for “early exit” or psychological withdrawal to mitigate further exhaustion and economic precarity. This underscores that AI’s impact is not a monolithic outcome but is deeply mediated by the socio-economic vulnerabilities of the agent.

Furthermore, this study aligns with recent scholarly discourse suggesting that AI effects are not uniform but are contingent upon organizational settings and contextual job structures (Bankins et al., 2024). By demonstrating that employment intentions fluctuate across different regional contexts in China, this research extends the “non-uniformity” thesis into the geography of transition. Consistent with the perspective that technological adaptability in an aging workforce depends on a constellation of individual and societal support structures (Ranasinghe et al., 2024), our findings suggest that future theoretical models must elevate the role of professional background and nuanced work experience in AI-related employability analyses.

Finally, the study provides a broader methodological contribution by highlighting the “concept drift” inherent in digital labor markets. As AI capabilities and environmental conditions undergo rapid evolution, the stability of captured empirical relationships cannot be taken for granted. This necessitates a transition toward dynamic modeling, where analytical inferences and data reassessments are periodically updated to maintain reliability and accuracy (Hinder et al., 2024).

6. Conclusion

The study explores the impact of artificial intelligence (AI) technology on the employment intentions of individuals nearing retirement within the context of delayed retirement policies, with a focus on factors such as gender, education, region, and occupation type. Key findings include that males and those with higher educational attainment exhibit significantly stronger employment intentions compared to females and individuals with lower educational attainment. This trend is particularly pronounced in the eastern and northern regions, where AI technology is more widely adopted, leading to increased retirement pressure among lower-educated populations. Contrary to common belief, concerns about AI replacing jobs are more prominent in the eastern regions, yet employees in these areas are also more willing to improve their competitiveness through digital skills training. Additionally, the study finds that household economic pressure significantly influences employment intentions, with those facing financial burdens tending to retire on time or even earlier. These conclusions offer a unique perspective on the employment choices of different groups under delayed retirement policies.

Funding

No

Conflict of Interests

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

Reference

- [1] Banks, S., Ocampo, A. C. G., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, 45(2), 159–182.
- [2] Chang, C.-H., Xu, H., & Xie, B. (2023). Aging workforce in the context of technological advancements: Toward a socio-ecological model. *Work, Aging and Retirement*, 9(4), 323–328.
- [3] Dai, Y., Zhao, Z., Sui, J., & Xu, J. (2025). The impact of delayed retirement on labor employment, fertility rate and economic growth in China. *International Review of Economics & Finance*, 100, 104103.
- [4] Ferdous, S. (2023, June 14). Are older workers ready for an AI takeover at work? Oxford Institute of Population Ageing, University of Oxford.
- [5] Haan, P., & Tolan, S. (2019). Labor supply and fiscal effects of partial retirement: The role of entry age and the timing of pension benefits. *The Journal of the Economics of Ageing*, 14, 100187.
- [6] Hinder, F., Vaquet, V., & Hammer, B. (2024). One or two things we know about concept drift—a survey on monitoring in evolving environments. Part A: Detecting concept drift. *Frontiers in Artificial Intelligence*, 7, 1330257.
- [7] Ranasinghe, T., Grosse, E. H., Glock, C. H., & Jaber, M. Y. (2024). Never too late to learn: Unlocking the potential of aging workforce in manufacturing and service industries. *International Journal of Production Economics*, 270, 109193.
- [8] Truxillo, D. M., Brady, G., Fraccaroli, F., Zaniboni, S., & Yaldiz, L. M. (2026). Work is changing: Implications for an aging, age-diverse workforce. *Human Resource Management Review*, 36(2), 101131.
- [9] Úbeda-García, M., Marco-Lajara, B., Zaragoza-Sáez, P. C., & Poveda-Pareja, E. (2025). Artificial intelligence, knowledge and human resource management: A systematic literature review of theoretical tensions and strategic implications. *Journal of Innovation & Knowledge*, 10(6), 100809.
- [10] Valtonen, A., Piri, R., Kaasinen, E., & Raisamo, R. (2024). Mitigating employee resistance and achieving well-being in digital transformation with personalized assistance. *Information Systems Frontiers*. Advance online publication.
- [11] Wang, Z., Pan, S.-Y., & Dong, W. (2024). Public management approaches to an aging workforce: Organizational strategies for adaptability and efficiency. *Frontiers in Psychology*, 15, 1439271.
- [12] Xie, H., Fang, Y., & Zhang, Y. (2023). Providing digital technology training as a way to retain older workers: The importance of perceived usefulness and growth need. *Work, Aging and Retirement*, 9(4), 376–392.
- [13] Zhao, G., Zhou, D., & Fu, Y. (2024). Study on the impact of delayed retirement on the sustainability of the basic pension insurance fund for urban employees in China. *Sustainability*, 16(10), 3969.

Axial Compression Effect on Carbonation Resistance of Fly Ash Concrete: An Experimental and Modeling Study

Hu Li, Zhuang Tian*, Fulei Liu, Pengcheng Li

Heze Urban Construction Group, Shandong, 274000, China

*Corresponding author: Zhuang Tian, 17669328541@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: To investigate the durability of fly ash concrete under service conditions, this study systematically examined the effect of axial compressive load on its carbonation resistance. Prismatic specimens of ordinary Portland cement concrete (OPC) and concrete with 30% fly ash (FAC30) were prepared. Axial compressive loads at three levels (0%, 30%, and 60% of the ultimate strength) were applied using a self-designed long-term loading device. The specimens were then subjected to accelerated carbonation for 28 days in an environment with a CO₂ concentration of 20±1%, a temperature of 20±2°C, and a relative humidity of 70±3%. Carbonation depths were measured at 7, 14, and 28 days, and the distributions of internal pH value and calcium carbonate content were analyzed. The results showed that the axial load had a dual effect on the carbonation resistance of concrete. Under a low load level (30%), the carbonation depth of FAC30 decreased by approximately 6.5% compared to that of the unloaded specimens, indicating a certain inhibitory effect. However, under a high load level (60%), the carbonation depths of OPC and FAC30 increased significantly by 44.4% and 43.2%, respectively. The incorporation of fly ash substantially reduced the alkali reserve of concrete, resulting in a carbonation depth much greater than that of OPC, which was the dominant factor affecting carbonation resistance. Based on Fick's law and the experimental data, a prediction model for carbonation depth considering the axial load ratio was established as $X_c = X_0(1 + kS^m)$. Validation showed that the predicted values agreed well with the experimental results ($R^2 > 0.96$). This study provides an important theoretical and experimental basis for accurately predicting the service life of fly ash concrete structures under load.

Keywords: Fly Ash Concrete; Axial Compressive Load; Anti-Carbonation Performance; Carbonation Depth; Durability

Published: Mar 15, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1136>

1. Introduction

Carbonation of concrete is one of the key factors leading to durability deterioration, steel reinforcement corrosion, and even structural failure^[1]. Against the background of the carbon peaking and carbon neutrality goals, the use of fly ash, an industrial by-product, to partially replace cement has become an important technical approach for producing green high-performance concrete^[2]. The morphological effect, micro-aggregate effect, and pozzolanic effect of fly ash can effectively improve the workability of concrete, reduce hydration heat, and optimize the pore structure in the later stage^[3]. However, the pozzolanic reaction consumes Ca(OH)₂ produced by cement hydration, which reduces the alkali reserve inside the concrete and weakens its ability to resist CO₂ erosion, thereby accelerating the carbonation process^[4].

In practical engineering, concrete structural members such as columns and walls are subjected to long-term stress. Loading can

alter the internal microcracks and pore structure of concrete, thereby affecting the transport rate of CO_2 [5]. Currently, research findings on the coupling effect of loading and carbonation remain inconsistent. Most studies [6-7] indicate that tensile stress accelerates carbonation, while the effect of compressive stress is more complex. Long-term experiments conducted by Castel et al. [8] showed that moderate compressive load within the serviceability range slightly inhibits carbonation. Zhao et al. [9] reported that under long-term loading, the carbonation rate of fly ash concrete increases with the development of damage. Therefore, this study aims to simulate the real stress state of concrete structures by conducting accelerated carbonation tests on concrete with different fly ash contents under long-term axial compressive load. By systematically analyzing the evolution of carbonation depth, internal pH value, and calcium carbonate content, the coupling mechanism between loading and fly ash content is revealed. Finally, a prediction model for carbonation depth that can quantify the effect of axial compressive load is established, providing theoretical support for the durability design and service life prediction of fly ash concrete structures.

2. Methods

2.1 Materials and Experimental Design

The materials used in this study included cement, fly ash, aggregates, and a superplasticizer. The cement was P·I 42.5 grade Portland cement (Shandong Shanshui, China), and its physical properties are presented in Table 1. The fly ash was Class I fly ash (Borun, Gongyi, China), with silicon dioxide (SiO_2) and aluminum oxide (Al_2O_3) as its main chemical components, accounting for 45.1% and 24.2%, respectively. Natural river sand with a fineness modulus of 2.3 was used as the fine aggregate, and crushed stone with a continuous grading of 5–20 mm was used as the coarse aggregate. A polycarboxylate-based superplasticizer with a water-reducing rate of 40% was employed to adjust the workability of the mixtures. Concrete mixtures with a target strength grade of C30 were designed. The water-to-binder ratio was fixed at 0.6 for all mixtures. The detailed mix proportions are provided in Table 2.

Table 1. Physical properties of P·I 42.5 grade Portland cement

Specific surface area (m^2/kg)	Setting time (min)	Compressive strength (MPa)	Flexural strength (MPa)
335	Initial:170; Final:215	3d: 28.1, 28d: 53.0	3d: 6.4, 28d: 9.0

Table 2. Concrete mix ratio (kg/m^3)

Type	Concrete	Coal gangue powder	Water	Coarse aggregate	Fine aggregate	Water reducer
OPC	330	0	198	1162	719	0.74
FAC30	231	99	198	1160	719	0.74

2.2 Sample preparation and preservation

Prismatic samples with dimensions of 100 mm × 100 mm × 300 mm were cast for carbonation and axial compressive strength tests. Cubic samples with dimensions of 100 mm × 100 mm × 100 mm were also prepared to determine the standard compressive strength. All samples were compacted by vibration during molding and demolded after being covered with plastic film for 24 hours. They were then cured in a standard curing room at a temperature of $20 \pm 2^\circ\text{C}$ and a relative humidity of over 95% for 28 days.

2.3 Experimental Procedures

Mechanical Property Tests: The cubic compressive strength was determined using a YAW-3000D microcomputer-controlled compression testing machine. The axial compressive strength of prismatic samples was measured using a YE-200A hydraulic material testing machine to provide reference values for subsequent loading.

Long-term Axial Compressive Load Application: A self-designed spring-bolt loading system (Figure 1) was used to apply long-term loads corresponding to 30% and 60% of the axial compressive strength to the cured prismatic samples. Load variations were monitored in real time using strain gauges attached to the bolts and a data acquisition instrument. When the load loss exceeded 10%, manual compensation was performed to ensure the stability of the load level.

Accelerated Carbonation Test: The samples, together with the loading devices, were placed vertically into a TH-B type carbonation chamber. The chamber environment was controlled as follows: CO_2 concentration of $20 \pm 1\%$, temperature of 20

$\pm 2^{\circ}\text{C}$, and relative humidity of $70 \pm 3\%$. The carbonation ages were 7, 14, and 28 days.

Carbonation Depth Measurement: At each specified age, samples were removed and split using a compression testing machine. A 1% phenolphthalein ethanol solution was immediately sprayed onto the split surface. After the color stabilized (the non-carbonated area appeared purple-red, while the carbonated area remained unchanged), carbonation depths were measured using a vernier caliper at six selected points on each surface. The average value was taken as the carbonation depth of the sample.

Microscopic Performance Analysis: For samples after 14 days of carbonation, powder was collected layer by layer from the surface inward. The calcium carbonate content in each layer was measured using a GMH3151 digital pressure measuring instrument. Meanwhile, the pH value of the pore solution was determined using precision pH test paper.

Figure 1: Diagram and physical image of the self-made long-term load device.



3. Results and Discussion

3.1 Mechanical Property Analysis

The mechanical property test results (Table 3) showed that the 28d cubic compressive strength (21.3 MPa) and prismatic axial compressive strength (10.94 MPa) of FAC30 were significantly lower than those of OPC (25.9 MPa and 28.79 MPa, respectively). This was mainly attributed to the low pozzolanic activity of fly ash at an early age, resulting in slow early strength development. These results provided an accurate load ratio basis for subsequent long-term load application.

Table 3 Mechanical properties of concrete

Type	Cubic compressive strength (MPa)	Prismatic axial compressive strength (MPa)	30% Load Value (kN)	60% Load Value (kN)
OPC	25.9	28.79	288	576
FAC30	21.3	10.94	109.4	218.8

3.2 Carbonation Depth Analysis

3.2.1 Effects of Load and Fly Ash Content

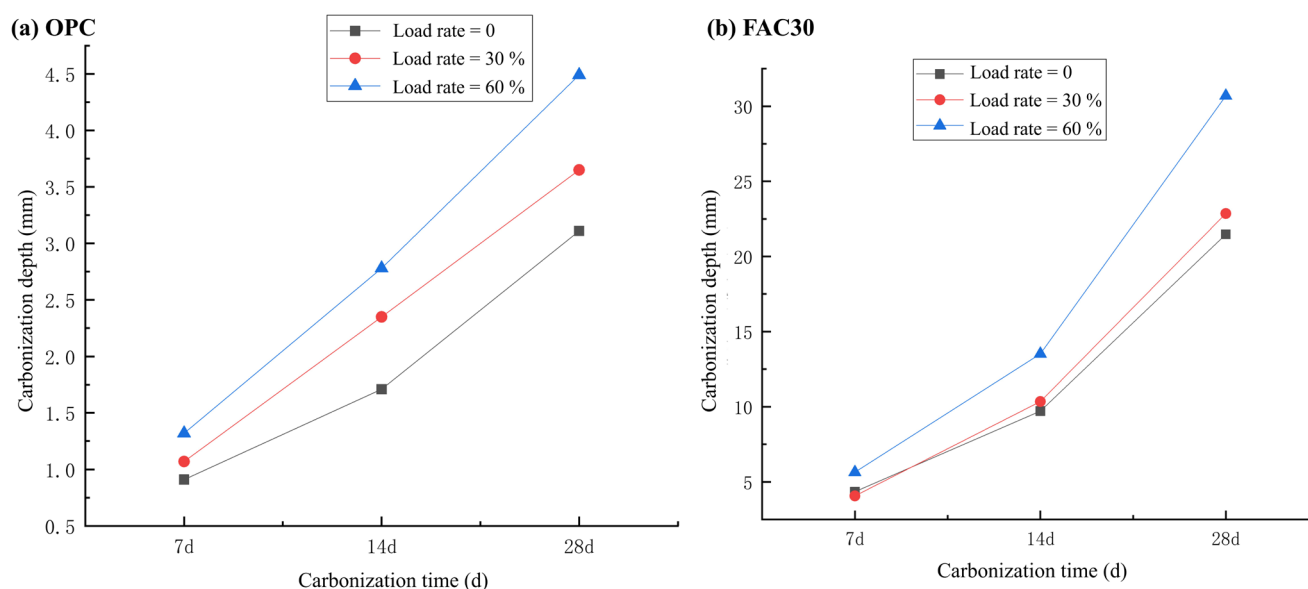
The carbonation depth test results under different conditions are presented in Table 4 and Figure 2. Under all load levels, the carbonation depth of FAC30 was substantially greater than that of OPC. For example, at 28 days, the carbonation depth of unloaded FAC30 (21.23 mm) was 6.8 times that of OPC (3.11 mm), confirming that the reduction in alkali reserve caused by fly ash consuming $\text{Ca}(\text{OH})_2$ is the primary factor responsible for the deterioration of carbonation resistance. Regarding the effect of load, a dual phenomenon was observed. Under a low load level of 30%, the carbonation depth of FAC30 at 7 days (4.06 mm) was slightly lower than that of the unloaded specimens (4.34 mm), suggesting that moderate compressive stress can compress initial microcracks and defects, temporarily densify the structure and hinder CO_2 diffusion. In contrast, when the load increased to 60%, the carbonation depth of both concrete types increased significantly. At 28 days, the carbonation depths of OPC and FAC30 increased by 44.4% and 43.2%, respectively, compared to their unloaded counterparts. This acceleration is attributed to the formation of new microcracks and the propagation and interconnection of existing cracks under high sustained load, which provide convenient pathways for rapid CO_2 intrusion.

Table 4 Carbonation depth (mm) of concrete under different load levels and time

Type	Load level	7d	14d	28d
OPC	0%	0.91	1.71	3.11
	30%	1.07 (+17.6%)	2.35 (+37.4%)	3.65 (+17.4%)
	60%	1.32 (+45.1%)	2.78 (+62.6%)	4.49 (+44.4%)
FAC30	0%	4.34	9.71	21.23
	30%	4.06 (-6.5%)	10.34 (+6.5%)	22.57 (+6.3%)
	60%	5.64 (+29.9%)	13.52 (+39.2%)	30.41 (+43.2%)

Note: Values in parentheses represent the change rate compared to the 0% load specimens in the same group.

Figure 2: Carbonation depth of concrete under different load levels over time.

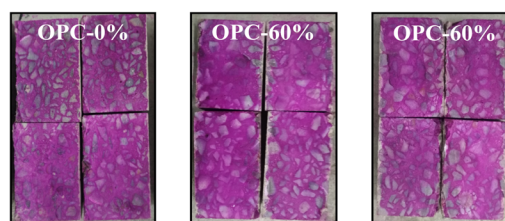


3.2.2 Carbonation Morphology Analysis

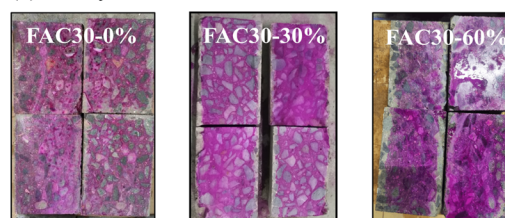
Figure 3 shows the phenolphthalein color development on the split surfaces of selected specimens after 28d of carbonation. It can be intuitively observed that the carbonated area (colorless) of FAC30 was much larger than that of OPC. Meanwhile, the carbonation front of specimens under high load (60%) was straighter and deeper, while that of unloaded or low-load specimens was relatively irregular, reflecting the “channeling” effect of load on CO₂ transport pathways.

Figure 3: Split surface morphology of concrete after 28d carbonation.

(a) 28d ordinary concrete



(b) 28d fly ash concrete

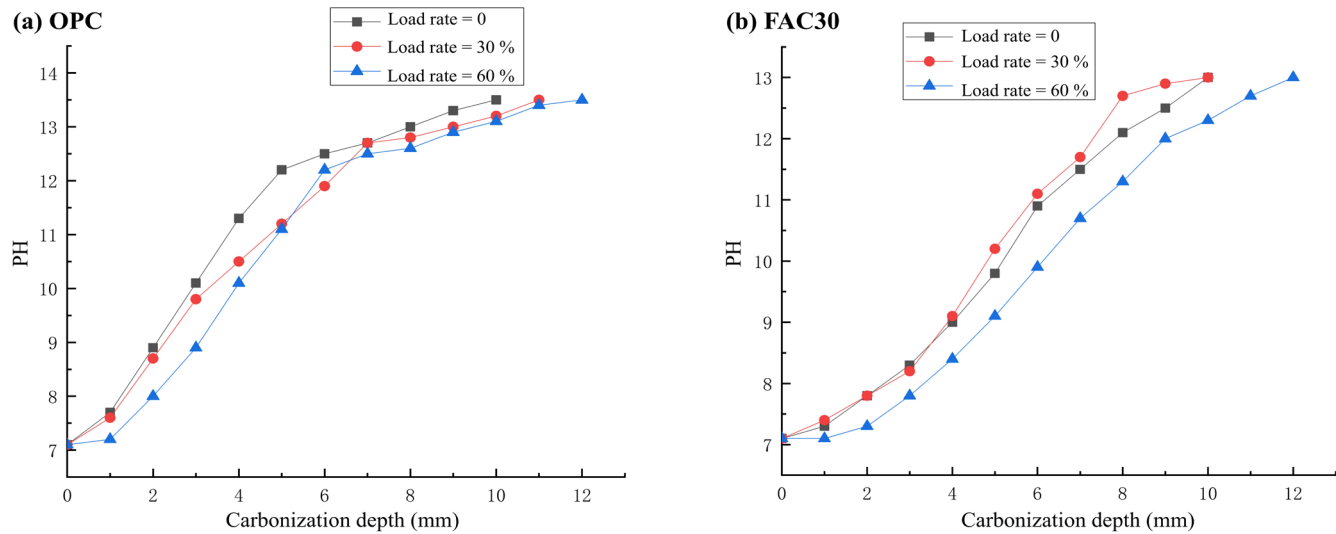


3.3 Internal pH Value and Calcium Carbonate Distribution

3.3.1 pH Value Distribution

Figure 4 shows the variation of internal pH value with depth after 14d of carbonation. It can be observed that: The pH value of FAC30 was lower than that of OPC at all depths, once again verifying its insufficient alkali reserve. For OPC, the higher the load, the lower the pH value at the same depth, indicating a deeper degree of carbonation. For FAC30, the pH value under 30% load within 4 mm depth was slightly higher than that of unloaded specimens, which corroborated the “inhibitory” phenomenon observed in the carbonation depth data. In contrast, 60% load caused a sharp decrease in pH value throughout the entire measured depth.

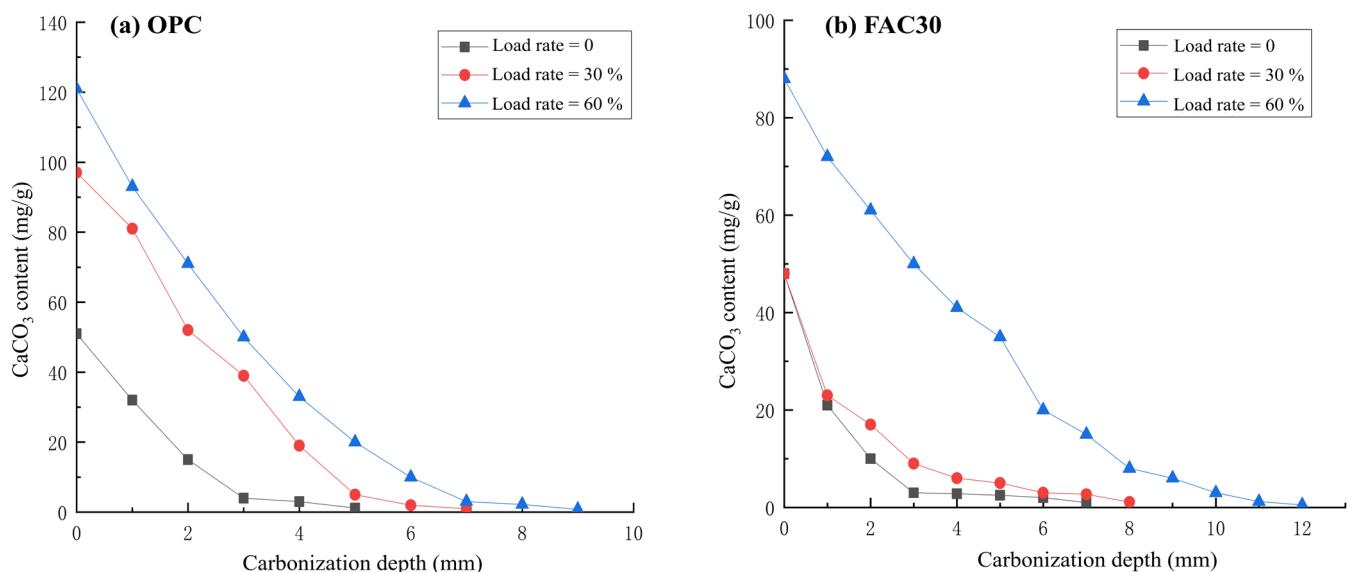
Figure 4: Internal pH value distribution with depth after 14d carbonation.



3.3.2 Calcium Carbonate Content Distribution

The calcium carbonate content distribution was obtained through layer-by-layer testing of specimens after 14d of carbonation (Figure 5). The calcium carbonate content was highest in the completely carbonated zone and gradually decreased with depth toward the non-carbonated zone. The CaCO_3 content in each layer of OPC was generally higher than that of FAC30, because OPC contained more carbonatable substances (Ca(OH)_2). The increase in load, especially 60% load, led to a higher CaCO_3 content at the same depth, indicating that more CO_2 had diffused to that location and participated in the reaction. This intuitively reflected the accelerating effect of load on the carbonation process.

Figure 5: Calcium carbonate content distribution after 14d carbonation.



3.3 Internal pH Value and Calcium Carbonate Distribution

3.4 Carbonation Depth Prediction Model

Based on Fick's first law, carbonation depth is generally proportional to the square root of carbonation time, namely:

$$X_c = K\sqrt{t} \quad (1)$$

In this study, the axial compressive load ratio S (S = Applied load/ultimate axial compressive load) was introduced as an influencing coefficient, and the following modified model was proposed:

$$X_c = X_0(1 + kS^m) \quad (2)$$

where, X_0 is the carbonation depth under unloaded condition, which can be fitted by $X_0 = Kt^a$, k and m are model parameters related to the material properties of concrete.

Fitting was performed using the 28d carbonation depth experimental data, yielding:

$$\text{For OPC: } X_0 = 0.15t^{0.91}, k = 0.18, m = 1.36$$

$$\text{For FAC30: } X_0 = 0.45t^{1.16}, k = 1.62, m = 2.6$$

For fly ash concrete at a carbonation age of 14d: when $S = 0.3$, $X_c = 10.34$; when $S = 0.6$, $X_c = 13.25$. Substituting these into the formula gave $k = 1.62$ and $m = 2.6$. The theoretical carbonation depth at 28d under 60% load ratio was calculated to be 30.69 mm, which was close to the experimental result of 30.41 mm, with an error of 1%. Therefore, this model could be applied to this experiment. The carbonation model for fly ash concrete under these conditions was $X_c = X_0(1 + 1.62S^{2.6})$. Similarly, the carbonation model for ordinary concrete under the same conditions was $X_c = X_0(1 + 0.18S^{1.36})$.

4. Rebounding and Defensive Comprehensiveness

This study investigated the effect of axial compressive load on the carbonation resistance of fly ash concrete. The main conclusions are as follows: Axial compressive load exhibited a significant dual effect on the carbonation resistance of fly ash concrete. A low load level of 30% slightly inhibited carbonation through compaction, while a high load level of 60% significantly accelerated the carbonation process by inducing microcracks. At 28 days, the carbonation depth increased by more than 43% under high load. Fly ash content was the dominant factor affecting the carbonation resistance of concrete. Under the same conditions, the carbonation depth of concrete with 30% fly ash replacement was more than six times that of ordinary concrete. This was mainly attributed to the consumption of alkaline substances by the pozzolanic reaction. The underlying mechanism of the load effect was revealed. Load regulated the CO_2 diffusion rate by altering the pore structure and crack network of concrete. Low load optimized the structure, whereas high load damaged it. A carbonation depth prediction model was established, incorporating the effects of time, fly ash content, and axial compressive load ratio. The model showed high predictive accuracy and can serve as a quantitative tool for durability design and service life assessment of engineering structures. When applying fly ash concrete in practice, appropriate measures such as reducing the water-to-binder ratio, adding anti-carbonation admixtures, or applying surface protection are recommended to compensate for its limited carbonation resistance, especially in members subjected to high stress levels.

Funding

No

Conflict of Interests

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

Reference

- [1] Papadakis, V. G., Vayenas, C. G., & Fardis, M. N. (1991). Fundamental modeling and experimental investigation of concrete carbonation. *ACI Materials Journal*, 88(4), 363-373. <https://doi.org/10.14359/1863>
- [2] Zhang, Y., Lu, L., & Yao, Y. (2016). Research on carbonation effect of fly ash concrete. *Construction Technology*, 45(S2), 526-529.
- [3] Liu, J. (2016). Research on the carbonation characteristics of concrete of different strength grades. *China Concrete and Cement Products*, (6), 24-28. <https://doi.org/10.19761/j.1000-4637.2016.06.006>

- [4] Nath, P., & Sarker, P. (2011). Effect of fly ash on the durability properties of high strength concrete. *Procedia Engineering*, 14, 1149-1156. <https://doi.org/10.1016/j.proeng.2011.07.144>
- [5] Zhao, Q., He, X., & Zhang, J. (2017). Effect of long-term loading on carbonation resistance of fly ash concrete. *Journal of The Chinese Ceramic Society*, 45(2), 86-91. <https://doi.org/10.14062/j.issn.0454-5648.2017.02.12>
- [6] Tu, Y., & Lu, Z. (2003). Experiment and research of prestressed concrete structure in carbonation corrosive environments. *Journal of Southeast University (Natural Science Edition)*, 33(5), 573-576.
- [7] Chen, W., Xu, Y., & Geng, J. (2003). Effect of carbonization and stress on performance of concrete resistance to chloride ion penetration. *Bulletin of the Chinese Ceramic Society*, 33(5), 573-576. <https://doi.org/10.16552/j.cnki.issn1001-1625.2015.05.003>
- [8] Castel, A., François, R., & Arliguie, G. (1999). Effect of loading on carbonation penetration in reinforced concrete elements. *Cement and Concrete Research*, 29(8), 561-565. [https://doi.org/10.1016/S0008-8846\(99\)00017-4](https://doi.org/10.1016/S0008-8846(99)00017-4)
- [9] Hao, X., Zhang, W., & Liu, Y. (2017). The carbonation properties of the fly ash concrete under the axial compression load. *Journal of Agricultural University of Heibei*, 40(6), 108-112. <https://doi.org/10.13320/j.cnki.jauh.2017.0134>

Research on the Path of Carbon Market Risk Governance Empowered by Accounting Tools

Xinyao Zhang*

Master Candidate of Accounting, School of Economics and Management, Yan'an University, Yan'an, 716000, China

*Corresponding author: Xinyao Zhang, zhangxinyao1107@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 background of the implementation of China's "Dual Carbon" strategy and closely focusing on the practical needs of carbon market risk governance, this paper systematically explores the internal logic and practical paths of accounting tools empowering carbon market risk governance. The research first conducts risk identification and measurement analysis, and then clarifies the characteristics and impacts of four types of core risks in the carbon market: (1) Market manipulation and price fluctuation risks impact the market pricing mechanism and the financial stability of enterprises; (2) Data quality and integrity risks weaken the credibility of transactions and lead to regulatory failure; (3) Policy arbitrage and carbon leakage risks damage market fairness and hinder the achievement of emission reduction targets; (4) Quota allocation and market imbalance risks result in resource misallocation and inhibit market liquidity. On this basis, a multi-dimensional risk governance path system of accounting tools is constructed: fair value measurement, hedge accounting and other tools are used to stabilize price fluctuations; carbon accounting information systems and audit supervision are relied on to strengthen credit risk management; policy tracking and accounting adaptation are adopted to respond to policy change risks; risk transfer and sharing mechanisms are established by means of risk reserve accrual and carbon financial tool innovation. This research clarifies the core functions and application value of accounting tools in carbon market risk governance, provides theoretical support and practical reference for improving the carbon market risk governance mechanism and enhancing market operation efficiency, and helps the carbon market better play its role in resource allocation in the low-carbon transformation.

Keywords: Accounting Tools; Carbon Market; Risk Governance

Published: Mar 15, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1175>

1. Introduction

1.1 Research Background

The global climate governance process is accelerating, and the in-depth implementation of China's "Dual Carbon" strategic goals has made the carbon market a core hub for guiding social resources to be allocated to the low-carbon field. As the core carrier of carbon emission rights trading, the stable operation of the carbon market is directly related to the efficiency of achieving emission reduction targets. However, the superposition and exposure of multiple risks in the current market operation have significantly restricted its governance efficiency.

In terms of risk performance, market manipulation and violent price fluctuations damage the fairness of the carbon pricing mechanism, directly lead to the distortion of the fair value measurement of enterprises' carbon assets, and trigger fluctuations

in financial statements and deviations in business decisions; frequent problems of distorted data quality and lack of integrity make carbon emission accounting deviate from the accounting recognition and measurement standards, and there are loopholes in the verification link due to the lack of systematic accounting supervision tools, which weakens the market participants' trust in carbon trading data; the existence of policy arbitrage and carbon leakage risks leads to inconsistent calibers of carbon cost accounting among different regions and industries, violates the principle of comparability of accounting information, and causes unfair competition; unreasonable quota allocation is likely to trigger market supply and demand imbalances, resulting in quota surplus or shortage, which not only increases the risk of carbon asset impairment, but also inhibits the application efficiency of accounting tools in risk hedging due to insufficient liquidity of trading targets.

Against this background, carbon market risk governance is in urgent need of systematic support from professional accounting tools. Relying on its core advantages in the recognition and measurement of carbon assets and liabilities, risk quantitative measurement, and full-process supervision and accounting, accounting tools have become a key starting point for empowering carbon market risk governance. Traditional carbon market risk governance mostly relies on administrative supervision and spontaneous market regulation, lacking quantifiable accounting standards and dynamic monitoring mechanisms, with limitations such as lagging risk early warning and single control methods, making it difficult to cope with complex and changeable risk situations.

In contrast, accounting tools can accurately capture the impact of carbon price fluctuations on corporate finance through fair value measurement, realize the full-process traceability and audit supervision of emission data relying on the carbon accounting information system, and effectively prevent and control data fraud and credit risks; the hedging accounting rules can be used to realize the hedge accounting of carbon price risks, and a risk sharing mechanism can be constructed by means of risk reserve accrual and accounting treatment of carbon financial tools, helping market entities to cope with the impact of price fluctuations and policy changes steadily.

At present, academic and practical research on the systematic empowerment of carbon market risk governance by accounting tools is still scattered, and a theoretical model and practical path covering the whole chain of risk identification, measurement, control and sharing with accounting tools as the core has not been established. This leads to the lack of systematic guidance for the application of tools such as fair value measurement, carbon audit and hedge accounting, making it difficult to give full play to the core value of accounting tools in risk quantification, process supervision and information disclosure. This practical contradiction highlights the urgency and necessity of carrying out this research.

1.2 Research Significance

The theoretical significance and practical value of this research can be expanded from multiple dimensions: at the theoretical level, by constructing a correlation framework between accounting tools and carbon market risk governance, this research can inject an interdisciplinary research perspective into the carbon market risk governance theory and enrich the theoretical system in this field; at the same time, it can expand the application boundary of environmental management accounting theory, extending environmental management accounting from the traditional level of carbon information disclosure and cost accounting to the dynamic risk control scenario of the carbon market, and providing theoretical support for the innovative development of this theory in the field of low-carbon governance.

At the practical level, the application paths of accounting tools proposed in this research for the four types of core risks in the carbon market, such as stabilizing price fluctuations with fair value measurement and strengthening credit risk control relying on carbon audit, can provide market entities such as enterprises and financial institutions with directly implementable risk governance tools, help them reduce the impact of risks on business decisions and financial stability, and assist market entities in achieving stable operation in the low-carbon transformation; at the same time, it can provide a reference for regulatory authorities to optimize the carbon market governance mechanism, which not only helps regulatory authorities improve the accuracy and efficiency of risk early warning relying on accounting tools, but also provides practical basis for improving quota allocation rules and designing risk reserve systems.

More importantly, by giving play to the empowering role of accounting tools in carbon market risk governance, the restrictions of risks on the fairness of the carbon pricing mechanism and the standardization of market operation can be reduced,

guiding more social capital to flow to the low-carbon industry, and providing a solid institutional and tool support for the on-schedule achievement of the “Dual Carbon” strategic goals.

2. Literature Review

2.1 Carbon Trading Theory

Against the background of intensified global climate change, reducing greenhouse gas emissions is the common responsibility of all mankind. The Kyoto Protocol first introduced market mechanisms into greenhouse gas emission reduction, and since then, carbon emissions have been endowed with commodity attributes and can be freely traded in the market. Carbon trading refers to a transaction method based on greenhouse gas emissions^[1]. Carbon emission rights trading, a form of carbon trading, is realized by the government setting emission reduction targets, allocating carbon emission quotas and limiting emissions within a certain period.

Policies related to carbon emission rights trading mainly involve the issue of “carbon emission rights trading”, that is, the policy measures adopted by the government to regulate the total carbon emissions and the division and trading of carbon emission quotas^[2]. Since “carbon trading” has become a widely used term in academia, for the sake of concise expression of the concept, this paper refers to the carbon emission rights trading policy as the “carbon trading policy”. In 2020, China announced that its carbon dioxide emissions would peak before 2030 and strive to achieve carbon neutrality before 2060. The national carbon emission rights trading market is one of the core policy tools to achieve the goals of carbon peaking and carbon neutrality.

2.2 Environmental Management Accounting Theory

In the 1970s, with the in-depth development of theoretical research such as Demsetz’s “Research on the Transformation of Social Costs of Pollution Control” and Marin’s “Pollution Accounting Issues”, environmental management accounting has gradually attracted the attention of academia. Since the 21st century, research on environmental protection-related issues in the accounting field has increased day by day.

Compared with foreign countries, China has attached less importance to the research of environmental management accounting and started late. Through literature combing, China’s environmental management accounting research mainly focuses on three aspects: first, the research on the relationship between the environment and accounting, Wu Wen studied the relationship among cost stickiness, relational transactions and corporate environmental accounting information disclosure^[3]; second, the research on the synergistic effect of the government, enterprises, environmental organizations and other subjects in environmental management, Han Zhiqing and Han Ruixue put forward the Nash equilibrium game analysis of environmental governance in operational research^[4], indicating that solving the negative externalities of environmental problems needs to rely on external forces such as the government and play a joint role with enterprises and environmental protection organizations; third, the interdisciplinary research of environmental management accounting with environmental science, operational research, management science and other disciplines, scholars such as Zheng Qiong and Liu Chunying and Wang Qinglin integrated ESG into environmental accounting, providing an important guarantee for the sustainable development strategy of enterprises^{[5][6]}.

2.3 Risk Management Theory

The economic crisis in the 1930s prompted the emergence of risk management ideas. In order to cope with the huge impact, people began to pay attention to the risk management of various activities, so as to control or reduce losses as much as possible. In the mid-20th century, after Meyer and Hersh of the United States published “Enterprise Risk Management” in the Insurance Manual, research on risk management gradually became systematic and specialized, which made “risk management” gradually develop into a discipline, and enterprises entered the stage of traditional risk management.^[7]

Operators began to attach importance to risk management activities, and the main measures were to avoid and transfer risks. The representative theories of this stage include Markowitz’s mean-variance theory, the capital asset pricing model and option pricing model. However, the relevant research is relatively single and one-sided, lacking hierarchy and systematicness, and even lacking a strategic height. Until the 1980s, with the tremendous changes in the social environment and other factors, developed capital markets and tools have become an important force promoting the development of the modern risk

management stage. Enterprises pay more attention to the combined risks of various departments and their comprehensive impacts, and carry out risk management based on this to improve operational efficiency.

At the same time, since China resumed insurance business, research and application related to risk management have emerged. Zhang Xuchu and others believe that risk management is a broad concept, whose purpose is to suppress the random economic influencing factors and ensure that business activities are in an orderly state. Therefore, they propose that when exploring advanced foreign experiences, Chinese enterprises should analyze various risks faced by enterprises from an overall perspective, especially the control and management of dynamic risks.^[8]

At the beginning of the 21st century, the development of risk management theory entered the stage of enterprise risk management (ERM). In 2004, COSO upgraded the content of the Internal Control-Integrated Framework and launched the Enterprise Risk Management-Integrated Framework, transforming the internal control of enterprises from “financial report-oriented” to “risk management-oriented”^[9]. The core of the ERM theory comes from the concepts and methods of risk management in the report. In 2017, COSO further improved the framework content, and its scope of application was clearly expanded from enterprises to organizations of any type and scale. Based on this discovery, Wang Hui proposed that the framework can provide a good opportunity for embedding risk management in the establishment and improvement of internal control of administrative institutions in China, which is conducive to the connection and integration of the new framework of government internal control and risk management, forming a late-mover advantage.^[10]

3. Carbon Market Risk Identification and Measurement

3.1 Risk Identification Methods

3.1.1 Data Analysis Method

Carbon price data analysis should be carried out from a multi-dimensional perspective: first, time series analysis (such as ARIMA model) is used to identify price trends and seasonal laws, with a focus on abnormal price fluctuations before and after the compliance period; calculate historical volatility (such as 30-day annualized volatility), and then use GARCH family models to capture volatility clustering characteristics; at the same time, conduct correlation analysis, focusing on the linkage between carbon prices and energy prices (e.g., the strong negative correlation of -0.7 between EU carbon prices and TTF natural gas prices) and macroeconomic indicators (such as PMI index); it is also necessary to carry out monitoring of market liquidity indicators, including Z-score analysis of daily trading volume, moving average monitoring of bid-ask spreads, and changes in open interest, to identify liquidity hidden dangers caused by insufficient market depth.

3.1.2 Model Construction Method

Model construction should adopt a multi-level method: improved ARIMA models (such as SARIMA considering seasonal factors) combined with XGBoost feature selection can be used for short-term forecasting, with a forecasting accuracy of more than 85%; EGARCH (1,1) model is recommended for volatility modeling, which can effectively capture the asymmetric characteristics of “sharp rise and slow fall” of carbon prices; for medium and long-term analysis, a dynamic CGE model needs to be constructed, incorporating more than 50 parameters such as policy variables (e.g., quota reduction rate) and technology learning curves (e.g., photovoltaic cost decline); in terms of machine learning, LSTM neural network performs well in processing nonlinear relationships, but attention should be paid to preventing overfitting, and Bayesian optimization is recommended for hyperparameter tuning.

3.2 Risk Measurement Methods

In the context of the carbon market, commonly used risk measurement methods mainly include two categories: relative risk measurement methods and absolute risk measurement methods.

3.2.1 Relative Risk Measurement Method

The relative risk measurement method is mainly used to reflect the sensitivity relationship between the fluctuation of market factors and the price change of financial assets^[11]. Commonly used relative risk measurement indicators include Duration and Convexity. As a measurement indicator, Duration reflects the sensitivity of bond prices to interest rate changes. Although it is not common to directly use Duration to measure market risks in the carbon market, similar concepts can be used to analyze the sensitivity of carbon emission rights prices to policy changes, market supply and demand changes and other factors. By

calculating the Duration of carbon emission rights prices, the degree and direction of the impact of market factors on price changes can be grasped.

Convexity is an indicator to measure the nonlinear relationship between bond price changes and interest rate changes. In the carbon market system, Convexity can be used to analyze the nonlinear relationship between carbon emission rights price changes and market factor changes. When market factors change significantly, Convexity can reflect the acceleration or deceleration effect of price changes, thereby helping market participants to evaluate market risks more appropriately.

As an important way to assess risks, stress testing and scenario analysis play an extremely critical role in the carbon market. Stress testing focuses on evaluating the risk status of the market under extreme market conditions. Through carefully designed simulation scenarios, such as an immediate change in the policy environment or a sharp decline in market demand, it analyzes in detail the effects of these extreme scenarios on carbon emission rights prices and the overall market stability.

Scenario analysis focuses on the prospective setting and analysis of possible future market scenarios. It comprehensively considers multiple dimensions such as changes in the macroeconomic environment, forecasts of policy orientation and the trend of technological progress, aiming to compare the performance of market risks under different scenarios and provide richer and more comprehensive decision-making reference for market participants. These two methods jointly constitute the main part of risk measurement in the carbon market, helping investors and policymakers grasp market changes more accurately and formulate effective risk control strategies.

3.3 Methods for Quantitative Analysis of Core Market Risks

3.3.1 Financial Engineering Methods

Financial engineering methods combine finance theory with mathematics, statistics, computer science and other technologies to solve financial problems. In the quantitative analysis of market risks, commonly used financial engineering methods include VaR model, CVaR model and Copula model.

3.3.1.1 VaR Model

VaR (Value at Risk) refers to the maximum possible loss of the value of a financial asset or securities portfolio within a specific period in the future at a certain probability level (confidence level). The VaR model is widely used in financial risk management due to its simplicity and easy understanding. The calculation of the VaR model is usually based on historical simulation method, variance-covariance method or Monte Carlo simulation method.

In the carbon market, the VaR model can also be used to measure market risks. By setting an appropriate confidence level and holding period, the maximum possible loss value of the carbon emission rights portfolio in the future period can be calculated, thus providing important risk reference information for market participants^[12].

The mathematical expression of the VaR model is:

$$\text{Prob}(\Delta P > \text{VaR}) = \alpha$$

Among them, ΔP is the loss of the investment portfolio during the holding period, and α is the confidence level.

The calculation of the VaR model usually involves the following steps: Determine the asset portfolio:

clarify the carbon emission rights portfolio for which risks need to be measured; Select historical data: collect historical transaction data related to the carbon emission rights portfolio, which should include information such as price, trading volume and time;

Determine the confidence level and holding period: set an appropriate confidence level and holding period according to risk management objectives and market environment. The confidence level indicates the reliability of the calculation results, generally choosing 95% or 99%; the holding period represents the time range under investigation, which can be one day, one week, one month, etc;

Calculate the statistical distribution: use historical data to fit the statistical distribution of the returns of the carbon emission rights portfolio, usually choosing normal distribution, t-distribution or historical simulation method.

Calculate the VaR value: calculate the VaR value according to the selected confidence level and statistical distribution. Taking the normal distribution as an example, if the confidence level is set to 95%, the corresponding α value is 0.05, that is, the 5% right quantile under the normal distribution. The corresponding quantile z value can be obtained by looking up

the table, and then the VaR value is calculated by the formula $\text{VaR} = z \times \sigma \times W$, where σ is the standard deviation of the returns of the carbon emission rights portfolio, and W is the initial investment amount;

Result analysis and reporting: analyze the market risk level of the carbon emission rights portfolio according to the calculated VaR value, and write a risk report to provide risk disclosure and decision support to relevant parties.

3.3.1.2 CVaR Model

CVaR (Conditional Value at Risk) model refers to the average level of losses exceeding the VaR value at a certain confidence level. CVaR is a further improvement of VaR, which considers the size of the average loss exceeding the VaR level. In other words, CVaR gives the expected average loss that may be suffered when the VaR value has been determined and the market conditions are worse than the VaR setting. CVaR can capture extreme market risks better than VaR, so it is regarded as a more robust risk measurement indicator in some occasions.

The calculation of CVaR is relatively complex and usually needs to be realized through simulation methods or optimization technologies. Specifically, on the basis of generating a large number of market scenarios, all loss values exceeding the VaR threshold can be calculated, and the average of these loss values is taken as the estimated value of CVaR.

The mathematical expression of the CVaR model is:

$$\text{CVaR} = E[X|X > \text{VaR}]$$

Among them, (X) is the loss of the investment portfolio.

CVaR can be obtained by first calculating VaR and then calculating the average of losses exceeding the VaR part. The specific calculation method varies according to the adopted model. In the Monte Carlo simulation, after calculating the VaR, for all samples in the simulation results where the loss exceeds the VaR, their average loss is calculated, and this average loss is the value of CVaR.

3.3.1.3 Copula Model

The core of the Copula model is to “separate the marginal distribution from the correlation structure”, that is, first fit the return distribution (marginal distribution) of each asset in the portfolio separately, then describe the correlation (including linear and nonlinear correlation) between assets through the Copula function, and finally construct the joint distribution of the multi-asset portfolio to realize the accurate quantification of portfolio risks. It breaks through the assumption of “linear correlation between assets” in traditional methods and is more in line with the complexity of asset correlation in the actual market (such as the nonlinear linkage between carbon prices, coal prices and electricity prices).

The basic mathematical expression of the Copula model is:

For an n -dimensional random vector $X = (X_1, X_2, \dots, X_n)$, its joint distribution function $H(X)$ can be expressed as:

$$H(X_1, X_2, \dots, X_n) = C(F_1(X_1), F_2(X_2), \dots, F_n(X_n))$$

Among them: $F_i(x_i)$ is the marginal distribution function of X_i ($i = 1, 2, 3, \dots, n$);

$C:[0,1]^n \rightarrow [0,1]$ is the Copula function, which captures the dependence structure between variables.

The calculation of the Copula model usually involves the following steps:

Marginal distribution fitting: for the yields of each asset in the portfolio (such as EUA, CCER, coal futures), select appropriate marginal distributions (such as t-distribution, dynamic distribution fitted by GARCH model) respectively;

Copula function selection: select a suitable Copula function according to the correlation characteristics between assets;

Parameter estimation and testing: estimate the parameters of the Copula function by the maximum likelihood estimation method, and select the optimal Copula model by the AIC and BIC criteria;

Portfolio risk quantification: based on the joint distribution, calculate the VaR and CVaR of the portfolio, or simulate the risk-return curve under different asset weights to provide a basis for portfolio optimization.

4. Types of Carbon Market Risks

4.1 Market Manipulation and Price Fluctuation Risks

Market manipulation and price fluctuation risks directly impact the market pricing mechanism and corporate financial stability, and are one of the most prominent risk types in the operation of the carbon market. Manipulation behaviors in the carbon market are mainly manifested in large institutions or traders using their capital advantages, information advantages

and market dominant positions to artificially intervene in the supply and demand of carbon quotas through centralized trading, related transactions and other means, thereby manipulating the trend of carbon prices.

This behavior was particularly typical in the early stage of the European Union Emission Trading System (EU ETS). In 2006, the system experienced an extreme case where the carbon price plummeted from 30 euros to 1 euro due to the superposition of factors such as centralized selling by institutions and quota surplus, which directly led to the instant shrinkage of the value of a large number of corporate carbon assets, and some enterprises relying on carbon asset pledge financing fell into financial difficulties.

The core means of manipulators can be divided into two categories: first, “cornering the market”, that is, continuously buying a large number of carbon quotas and holding them for a long time to artificially create an atmosphere of tight market supply, push the carbon price to rise irrationally, and sell them centrally to make profits when the price reaches the expected level; second, “centralized selling”, that is, selling a huge amount of held quotas in a short time to form a market panic selling sentiment, suppress the carbon price rapidly, and then buy back to cover positions at a low price.

In addition, the wide application of high-frequency trading algorithms has further aggravated the short-term volatility of the market. Its millisecond-level trading speed and automated trading logic may amplify the impact of market sentiment without fundamental support, leading to abnormal fluctuations of “sharp rise and sharp fall” in carbon prices, making them deviate from the reasonable range reflecting the actual emission reduction costs of the industry.

4.2 Data Quality and Integrity Risks

Data quality and integrity risks weaken the credibility of carbon transactions and lead to regulatory failure, which are core risks threatening the sustainable operation of the carbon market. The accuracy of emission data is the premise for the effective operation of the carbon market. Key links such as carbon quota allocation, transaction pricing and compliance assessment all rely on true and complete emission data support. Once the data is distorted, the operation logic of the entire carbon market will have a fundamental deviation. However, from the practice of the global carbon market, data fraud problems have occurred frequently, which have become a prominent bottleneck restricting market efficiency.

Common data fraud means show diversified and hidden characteristics: at the enterprise level, they are mainly manifested in falsifying activity level data, tampering with emission factors, and manipulating monitoring equipment; in the early stage of China’s carbon market, some high-energy-consuming enterprises artificially reduced the accounting emissions by falsifying the calorific value of coal burning and concealing the consumption of purchased electricity, and obtained quota surplus in violation of regulations; in the voluntary carbon market, fraud behaviors of project developers are more concentrated, including falsifying the baseline emissions of emission reduction, exaggerating the emission reduction effect of projects, or registering and trading the same emission reduction across markets repeatedly, which seriously disrupts the market order.

4.3 Policy Arbitrage and Carbon Leakage Risks

Policy arbitrage and carbon leakage risks damage market fairness and hinder the achievement of emission reduction targets, which are typical systemic risks under the background of the coexistence of globalization and regionalization of the carbon market. The core function of the carbon market is to guide enterprises to reduce emissions through price signals. However, the differences in policy implementation and the limitations of regulatory scope provide opportunities for enterprises to evade emission reduction responsibilities, ultimately leading to the impairment of the fairness of the carbon market and the significant reduction of the overall emission reduction effect.

The manifestations of policy arbitrage are diversified, and the core can be divided into two categories: first, “carbon leakage”, that is, high-carbon enterprises transfer their production bases and high-emission links to regions with low carbon pricing and loose control standards to avoid strict carbon constraint policies, resulting in the offset of local emission reduction efforts by newly added emissions in other places. Relevant research data from the European Union shows that carbon leakage may lead to the invalidation of up to 25% of emission reductions in the region, forming an embarrassing situation of “emission reduction without consumption reduction”; second, “industry arbitrage”, enterprises take advantage of the policy differences between carbon market regulated industries and unregulated industries, strip high-emission businesses to affiliated enterprises not included in the regulation, or transfer emission responsibilities through upstream and downstream transactions in the

industrial chain, so that their own quota compliance pressure is reduced, while the total carbon emissions do not decrease. In addition, there is “time period arbitrage”, that is, enterprises take advantage of the rule loopholes in the policy transition period and emission reduction target adjustment period to delay emission reduction investment or suddenly increase emissions to obtain short-term quota benefits.

4.4 Quota Allocation and Market Imbalance Risks

Quota allocation and market imbalance risks lead to resource misallocation and inhibit market liquidity, which are core structural risks affecting the play of the carbon market’s resource allocation function. As the core trading target of the carbon market, the rationality of quota allocation directly determines the market supply and demand balance and the effectiveness of price signals. Once the allocation mechanism has defects, it will trigger a chain reaction and hinder the healthy development of the carbon market.

Market imbalances caused by improper quota allocation are mainly reflected in two dimensions: first, “total amount imbalance”, that is, the total amount of quotas does not match the actual emission demand. If the proportion of free allocation is too high and the total amount setting is loose, it will cause a serious surplus of quotas. At this time, the carbon price signal is severely suppressed, unable to form an effective emission reduction incentive, and low-carbon technology investment and emission reduction behaviors are shelved due to the lack of income support; if the quota allocation is too tight and the total amount setting is strict, it will greatly increase the compliance costs of enterprises, especially for enterprises in traditional industries with high energy consumption and low profits, which may exceed their bearing capacity, trigger operational pressure and even industry recession risks, and instead hinder the low-carbon transformation of the industry.

Second, “structural imbalance”, that is, the uneven distribution of quotas among different industries and enterprises. If the differentiated factors such as enterprises’ emission reduction potential, technical level and production scale are not fully considered, and a “one-size-fits-all” allocation standard is adopted, it will lead to inefficient enterprises obtaining excess quotas for idle waste, while efficient emission reduction enterprises face quota shortages and need additional procurement. This unfair allocation directly causes the misallocation of carbon emission rights resources, violating the core original intention of the carbon market to optimize resource allocation.

The “stock effect” caused by quota allocation further aggravates market imbalance: some enterprises hoard unused quotas for a long time based on the prediction of future carbon price increases or to avoid compliance risks, resulting in a large number of quotas stranded in enterprise accounts and difficult to circulate in the market. This behavior directly reduces the effective market supply, suppresses trading activity and liquidity, makes it difficult for the carbon market to form a continuous and fair price marker, and thus affects enterprises’ emission reduction decision-making and the actual efficiency of carbon asset operation.

5. Carbon Market Risk Governance Paths under Accounting Tools

5.1 Price Fluctuation Risk Governance

5.1.1 Improve the Market Mechanism

Enhance market transparency. For example, the EU ETS regularly publishes carbon emission data, transaction information and quota supply and demand forecasts to alleviate information asymmetry, provide clearer decision-making basis for market participants, and reduce irrational trading behaviors caused by information deviation. Enriching trading varieties is the core point: introducing derivative transactions such as futures and options to provide basic tools for hedging price risks; the standardized application of fair value measurement, hedge accounting and other tools can further reduce the impact of price fluctuations on corporate finance-fair value measurement can timely reflect the specific changes in the value of carbon assets and support enterprises to accurately grasp risk exposure; hedge accounting enables qualified derivative transactions to achieve risk hedging with carbon assets, maintain the stability of corporate profits, and reduce financial fluctuations caused by price shocks.

5.1.2 Strengthen Policy Guidance and Supervision

The government needs to set long-term and clear carbon emission reduction targets and phased promotion plans to reduce market expectation disorders caused by frequent policy changes, provide policy support to stabilize carbon prices, and form

a sound regulatory pattern. For example, the Supreme People's Court issued guiding opinions on the trial of carbon market-related cases, clarifying the identification standards and penalty basis for illegal acts such as market manipulation and insider trading, and severely punishing acts that intentionally affect carbon prices; build a price fluctuation early warning mechanism, and implement temporary regulatory measures when the carbon price deviates from the reasonable range to the preset critical value to ensure the fairness, impartiality and information transparency of the market.

5.1.3 Strengthen Policy Guidance and Supervision

Market participants should use technical means such as big data and artificial intelligence to integrate multi-dimensional information such as policy trends, industry emission data and market transaction trends, build carbon price prediction models, and improve the accuracy and timeliness of decision-making. Establishing a sound risk management mechanism is also crucial: enterprises should integrate carbon price risks into the overall risk management system, evaluate the size of risk exposure combined with fair value measurement results, and adopt diversified investment, hedging and other ways to reduce risks caused by single price fluctuations.

5.1.4 Adopt Financial Tools for Risk Hedging

Market participants can lock in costs and obtain benefits by buying or selling futures contracts or option contracts linked to carbon emission rights, thereby reducing the impact of price fluctuation risks; from the perspective of enterprises, they can match and account for derivative transactions with their own carbon asset risks in accordance with the corresponding norms of hedge accounting to ensure that the hedging effect is accurately reflected in the financial statements. In the field of carbon asset management and investment portfolio optimization, enterprises should regard carbon emission rights as a core asset and carry out full-life cycle management practices, including procurement, holding, transaction and exit stages. Based on changes in fair value, adjust asset allocation strategies in a timely manner, and realize risk diversification and reduction by building a diversified investment portfolio with carbon assets and related derivatives of different risk levels.

5.2 Credit Risk Management

5.2.1 Unify Monitoring, Accounting and Reporting Standards

Formulate a national unified and transparent MRV standard system, refine the monitoring scope, data collection caliber, accounting methods and report formats of various industries and emission sources, and ensure the consistency and comparability of data accounting among different enterprises and regions. At the same time, clarify the selection rules and dynamic update mechanism of emission factors, and regularly revise the emission factor database in combination with the progress of industrial technology and the upgrading of environmental protection standards to reduce the fraud space caused by the flexibility of artificial selection and ensure the accuracy of data from the source.

5.2.2 Strengthen the Independence and Professionalism of the Verification Mechanism

Implement a qualification certification and dynamic supervision system for verification institutions, strictly select third-party institutions with professional capabilities to participate in verification work, and establish a rotation mechanism for verification institutions to prohibit the same institution from providing continuous verification services for the same enterprise, so as to avoid verification inaccuracy caused by interest correlation from the system. Establish a cross-review and random spot check mechanism for verification results, and the regulatory authorities or independent third-party audit institutions shall conduct sample verification of the completed verification reports, focusing on verifying the integrity of data collection and the compliance of accounting methods; for doubts found in the verification process, require enterprises to supplement explanations and provide supporting materials to effectively improve the credibility of verification conclusions.

5.2.3 Establish a Strict Legal Liability Investigation System

Construct a stepped punishment system of "administrative punishment + civil compensation + criminal liability", and clarify the legal boundary and liability consequences of data fraud behaviors: for enterprises that commit data fraud, in addition to imposing heavy fines and recovering illegally obtained quotas, include them in the list of dishonest enterprises in the national credit information sharing platform, and restrict their carbon market transaction rights and relevant policy support qualifications; for those that cause losses to other market entities due to data fraud, order them to bear civil compensation liabilities; if the circumstances are serious and suspected of a crime, pursue the criminal liability of the relevant responsible

persons in accordance with the law.

At the same time, clarify the joint compensation liability of verification institutions. If false data flows into the market due to verification negligence, the verification institutions shall bear corresponding compensation liabilities, forcing them to standardize their practice processes and adhere to professional bottom lines.

5.3 Response to Policy Change Risks

5.3.1 Strengthen Policy Research and Prediction

Enterprises need to set up a professional policy research team, which is specially engaged in closely following the domestic and foreign policy trends. The team should use a variety of channels and methods to conduct an all-round and in-depth analysis of the policy trend, accurately predict the possible impact of policy adjustments on the carbon market, and actively establish close cooperative relations with government departments, industry associations and research institutions to obtain the most timely and accurate first-hand policy information and authoritative interpretations.

The policy research team shall regularly conduct a scientific assessment of the impact of policy changes on enterprises' carbon emission management, cost control and market competitiveness, comprehensively predict various risks and challenges that may be brought by policy changes by adopting professional means such as simulation analysis and risk assessment, and put forward targeted and operable response measures accordingly^[13]. If a policy that may affect carbon emission costs is predicted to be introduced soon, formulate cost control measures or adjust production methods in advance.

5.3.2 Improve the Risk Early Warning and Response Mechanism

Enterprises should use modern advanced information technology to create an efficient risk early warning platform. The system can monitor policy trends and market changes in real time, timely identify potential policy change hidden dangers, and ensure that enterprises can quickly take feasible measures to respond before risks break out by scientifically formulating early warning indicators, reasonably selecting thresholds and building an effective response mechanism.

For various types of policy change risks, enterprises should formulate detailed and comprehensive emergency plans, clarify the structure of the emergency organization system, the division of responsibilities of each member and the clear processing process and other key contents. In addition, through regular simulation drills and emergency training, continuously improve the enterprise's emergency response capacity and efficiency, and ensure that when risks really come, they can be handled quickly and in an orderly manner to minimize losses.

5.3.3 Improve Internal Management and Construction

Enterprises should establish a sound institutional system related to carbon emission management, clearly present the carbon emission rights management process, carefully divide the responsibilities of each party, and establish a strict supervision and assessment mechanism. By strengthening internal control management, the enterprise's carbon emission activities are strictly in accordance with the relevant laws, regulations and policy requirements, standardized and legal, and the risks caused by policy changes are truly reduced.

Pay attention to improving the quality and ability of employees, enhance their in-depth understanding and proficient application of carbon market policies, and improve their sensitivity and response level to policies. Use various means such as organizational training and interactive exchange activities to comprehensively enhance the professional quality and comprehensive capabilities of employees, and build a solid talent support system for enterprises to successfully resist policy change risks.

5.3.4 Enhance International Exchange and Cooperation

Enterprises should actively participate in the cooperative exchange activities of the international carbon market, keep a close eye on the policy trends and development trends of the international carbon market, form good cooperative relations with international counterparts, jointly study effective ways to deal with policy change risks and market challenges, and actively learn from the rich experience and mature measures of advanced international countries in carbon market policy formulation, regulatory mechanisms and risk management.

In combination with the national conditions and the actual situation of the enterprise itself, continuously improve and optimize its own carbon emission management and risk management system, and enhance the competitiveness and

adaptability of the enterprise in the global carbon market. At the same time, enterprises can pay attention to the latest research results in the field of international carbon market research, such as the research on carbon trading pricing based on the B-S pricing model by Zhu Yuezhao, Chen Hongxi and Zhao Zhimin, understand the advanced international theories and methods on carbon emission rights pricing^[14], and provide a reference for enterprises to reasonably price and reduce risks in international carbon market transactions^[15].

5.4 Improvement of Quota Governance Efficiency

5.4.1 Establish a Scientific Baseline Allocation Method

Taking the advanced emission level of the industry as the core benchmark, combined with differentiated factors such as the actual production scale, product structure, energy type and emission reduction potential of enterprises, a mixed allocation mode combining the “baseline method + historical intensity method” is adopted. For industries with mature technologies and complete emission data, the baseline method is given priority, and quotas are accounted according to the emission benchmark value per unit product; for industries with a concentration of small and micro enterprises and large emission fluctuations, the historical intensity method can be appropriately retained and gradually transitioned to the baseline method.

At the same time, introduce an industry emission reduction coefficient, give additional quota rewards to enterprises that overfulfill the emission reduction targets, and reduce the subsequent quotas for enterprises that fail to meet the standards, so as to ensure that the quota allocation matches the enterprises’ emission reduction responsibilities and capabilities, guide the flow of carbon emission rights resources to efficient emission reduction enterprises, and reduce resource misallocation.

5.4.2 Implement a Dynamic Adjustment Mechanism for the Total Quota

Establish a normalized quota total assessment and adjustment mechanism, and flexibly revise the total quota every 2-3 compliance cycles in combination with factors such as the progress of national emission reduction targets, changes in industry production capacity and the effectiveness of energy structure transformation. Build a quota surplus recovery and shortage supplement channel: for the excess quotas idle by enterprises for a long time (such as unused for more than one compliance cycle), the regulatory authorities will recover them at the market fair price and auction them publicly; for enterprises with compliance difficulties affected by policy adjustments and sudden operational pressure, provide transitional quota lending or short-term supplementary support to avoid severe market fluctuations caused by short-term supply and demand imbalances.

In this process, realize risk transfer and sharing by means of risk reserve accrual and carbon financial tool innovation - require market entities to accrue risk reserves according to the scale of quota holdings or transaction volume to cope with sudden risks such as sharp fluctuations in quota prices and enterprise compliance defaults; encourage financial institutions to innovate products such as quota pledge financing, quota repurchase and carbon asset securitization to help enterprises activate stock quotas and diversify the risks of quota holding and compliance.

5.4.3 Set Quota Validity Period and Circulation Incentive Policies

Clarify the service life of quotas, stipulate that the validity period of unused quotas after compliance is 2-3 years, and they will be automatically cancelled after the expiration, so as to prevent enterprises from hoarding quotas for a long time and artificially creating a shortage of market supply from the system. At the same time, introduce targeted circulation incentive policies: reduce part of the transaction handling fees or give credit points for carbon market transactions to enterprises that take the initiative to put idle quotas into market transactions; establish a market maker system in the quota secondary market, encourage financial institutions to participate in bilateral quotation of quotas, and improve the activity of market transactions; support cross-regional quota circulation and replacement, break regional market segmentation, expand the scale of quota transactions, and further release market liquidity.

5.4.4 Strengthen the Supervision and Restraint of Quota Allocation and Use

Relying on the carbon market registration system and trading platform, build a full-process dynamic monitoring system for quota “allocation-holding-circulation-cancellation”, and track the flow of quotas and changes in positions in real time. Focus on monitoring abnormal behaviors such as centralized holding and frequent transfer of large amounts of quotas, and take disciplinary measures such as restricting subsequent quota allocation, levying quota idle taxes and fees, and public notice and criticism against enterprises that maliciously hoard quotas and manipulate market supply and demand; for enterprises that

defraud quotas by falsely declaring production data, recover the illegally obtained quotas and impose heavy fines to maintain a fair market order.

6. Conclusions and Prospects

6.1 Conclusions

Under the background of China's vigorous promotion of the "Dual Carbon" strategy, the demand for carbon market risk governance is becoming increasingly urgent. Based on this, this paper focuses on this demand and deeply explores the empowerment logic and specific practical paths of accounting tools in it, and draws the following conclusions:

The stable operation of the carbon market is restricted by four types of core risks. First, market manipulation and price fluctuations damage the fairness and impartiality of carbon pricing and cause sharp fluctuations in the value of carbon assets held by enterprises. Second, risks in data quality and integrity weaken the credibility of carbon market transactions and make it difficult for regulatory authorities to intervene effectively. Third, policy arbitrage and carbon leakage lead to unfair competition and offset the original emission reduction effects. Fourth, unreasonable quota allocation and imbalance lead to resource misallocation and thus inhibit market liquidity. These four types of risks are intertwined and jointly affect the normal play of the carbon market's resource allocation function.

Accounting tools can build a governance system covering the whole chain of carbon market risks. Specifically, fair value measurement and hedge accounting can be used to effectively stabilize carbon price fluctuations; carbon accounting information systems and audit supervision mechanisms can be relied on to strengthen the management and control of credit risks; policy tracking and accounting adaptation can be used to buffer the impact of policy changes; risk transfer can be realized by accruing risk reserves and innovating carbon financial tools.

On the whole, the empowerment value of accounting tools is reflected in many aspects. For market entities, it can provide feasible risk control schemes to help enterprises operate steadily on the road of low-carbon transformation. From the perspective of governance mode, accounting tools can promote the transformation and upgrading of the carbon market governance mode from "administrative-led" to "tool empowerment + regulatory coordination", and provide strong support for regulatory authorities to optimize relevant rules. At the theoretical research level, it can not only fill the research gap on the value of accounting tools in carbon market risk governance and enrich the content of interdisciplinary theories, but also expand the application boundary of environmental management accounting, laying a solid foundation for the standardized operation of the carbon market under the "Dual Carbon" goals.

6.2 Prospects

In the future, the research can be further explored from the following three directions: at the theoretical level, an interdisciplinary perspective can be introduced to deeply analyze the correlation and transmission paths between various risks in the carbon market, further clarify the functional differences of different accounting tools, and gradually build a more targeted risk governance toolbox, thus enriching the existing theoretical framework; in practical application, the combination of accounting tools and new technologies such as blockchain and artificial intelligence should be promoted to continuously optimize risk measurement and control methods. At the same time, the industry characteristics and regional differences should be considered to formulate more targeted implementation plans and enhance the applicability of the tools; in terms of policy coordination, it is necessary to explore the connection mechanism between accounting tools and carbon market regulatory policies, make it adapt to the development of international rules such as the Carbon Border Adjustment Mechanism, and improve the accounting standards for carbon financial derivatives, so as to expand the application scope of accounting tools in risk governance and provide strong support for the high-quality development of the carbon market.

Funding

No

Conflict of Interests

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

Reference

- [1] Perdan, S., & Azapagic, A. (2011). Carbon trading: Current schemes and future developments. *Energy Policy*, 39(10), 6040–6054.
- [2] Zhang, D., Karplus, J. V., Cassisa, C., & Zhang, X. (2014). Emissions trading in China: Progress and prospects. *Energy Policy*, 70, 759–771.
- [3] Wu, W. (2024). Cost stickiness, relational transactions and corporate environmental accounting information disclosure. *Communication of Finance and Accounting*, (01), 50–54.
- [4] Han, Z. Q., & Han, R. X. (2018). A literature review of environmental management accounting research in China. *Finance and Accounting Monthly*, (21), 119–126.
- [5] Zheng, Q., & Liu, C. Y. (2021). Research on environmental management accounting under the ESG system – Taking Ping An Insurance (Group) Company of China, Ltd as an example. *Commercial Accounting*, (23), 26–29.
- [6] Wang, Q. L. (2025). Research on environmental management accounting embedded with ESG. *Communication of Finance and Accounting*, (03), 163–169.
- [7] Yan, F. H., Dang, X., & Yan, W. H. (2007). Review of the development course and trend of risk management. *Modernization of Management*, (02), 30–33.
- [8] Zhang, X. C., Wei, H. L., & Deng, D. S. (1985). A preliminary study on risk management. *Jiangnan Tribune*, (12), 19–23.
- [9] Huang, J. (2018). Research on the evolution and development of COSO internal control framework. *Auditing Monthly*, (04), 38–40.
- [10] Yang, Y. F. (2016). Application of enterprise internal control theory in social insurance fund management. *China Collective Economy*, (28), 41–42.
- [11] Chai, S. L., & Du, M. (2021). Carbon emission rights trading market risk measurement and management. China Financial and Economic Publishing House.
- [12] Wipplinger, E. (2007). Philippe Jorion: Value at risk – The new benchmark for managing financial risk. *Financial Markets and Portfolio Management*, 21(3), 397–398.
- [13] Zhao, L. X., & Hu, C. (2016). Research on the influencing factors of China's carbon emission rights trading price – An empirical analysis based on structural equation model. *Price Theory and Practice*, (07), 101–104.
- [14] Zhu, Y. Z., Chen, H. X., & Zhao, Z. M. (2013). Research on carbon trading pricing based on B-S pricing model. *Science & Technology Progress and Policy*, 30(05), 27–30.
- [15] Liu, C. Z., Pan, A. L., & Xie, D. Y. (2014). Policy suggestions for improving the carbon emission accounting system of Chinese enterprises. *Economic Review*, (11), 67–71.

The Impact of the Digital Economy on Municipal Air Quality: Insights from Spatial Analysis

Xinyan Yang¹, Lijun Yuan^{1,2*}

1. School of Business, Xinjiang Normal University, Urumqi, 830017, China

2. Research Center for High-Quality Industrial Development in the Core Area of the Silk Road Economic Belt, Urumqi, 830000, China

*Corresponding author: Lijun Yuan, 379713954@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: Air quality is a core element concerning public health and sustainable development. As the backbone of new quality productive forces, the digital economy is quietly emerging as a significant driver of air quality improvement. Based on prefectural-level city data in China from 2020–2023, and after confirming the spatial correlation between the digital economy and air quality, this paper first employs bivariate Moran's index to examine the spatial association characteristics between them. Subsequently, variance inflation factors are used to diagnose multicollinearity among annual variables. By constructing an ordinary least squares regression model, we investigate the overall impact effect. We then introduce a geographically weighted regression model to identify and estimate the spatial heterogeneity inherent in their relationship. Furthermore, the influence coefficient of the digital economy is utilized to analyze the spatial pattern and evolutionary trend of its impact on air quality. The results indicate a significant spatial dependence between the development of the digital economy and urban air quality, with this correlation pattern exhibiting a directional shift during the sample period. The geographically weighted regression model demonstrates superior goodness-of-fit and overall model adaptability compared to the ordinary least squares model, underscoring the robust spatial non-stationarity of the digital economy's impact on air quality. Further analysis reveals a clear East-West differentiation pattern in the spatial distribution of the digital economy's influence coefficients. This study not only provides spatial empirical support for exploring the complex relationship between the digital economy and environmental quality but also offers certain policy reference value for promoting regionally coordinated emission reductions and accelerating the green digitalization process.

Keywords: Digital Economy; Air Quality; Air Quality Index; Moran's Index; Geographically Weighted Regression Model

Published: Mar 20, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1184>

1. Introduction

Air quality constitutes a fundamental public resource for human survival and development, directly impacting global public health, sustainable development of economy and society, and regional ecological security. According to estimates by the World Health Organization (WHO), air pollution causes over two million premature deaths annually worldwide, highlighting its severity as a major environmental health risk factor^[1]. In China, the rapid economic and social development since the reform and opening-up has also accumulated a series of ecological and environmental issues, with air pollution being

particularly prominent, forming a critical constraint on high-quality development. According to the 2024 Report on the State of the Ecology and Environment in China, 117 out of 339 cities at prefecture level and above still failed to meet the air quality standards in 2024, accounting for 34.5% of the total, indicating that the task of improving air quality remains arduous^[2]. In response to this challenge, China has integrated air quality improvement into the core agenda of its national development strategy. In November 2023, the State Council issued the Action Plan for Continuous Air Quality Improvement, outlining a systematic approach to use sustained air quality improvement to drive high-quality economic development. By strengthening pollution source control and fostering green, low-carbon production and lifestyle, the plan strives to achieve a triple win of environmental, economic, and social benefits^[3]. Concurrently, the Proposal of the Central Committee of the Communist Party of China on Formulating the Fifteenth Five-Year Plan for National Economic and Social Development, adopted at the Fourth Plenary Session of the 20th CPC Central Committee, explicitly emphasized accelerating the comprehensive green transformation of economic and social development to build a Beautiful China^[4]. Furthermore, the coordinated development of the economy and environment remains a central topic in sustainability research. Studies indicate an overall upward trend in the coupling coordination level between China's regional economy and air quality^[5]. Simultaneously, promoting green development and improving air quality have been shown to help narrow the consumption gap between urban and rural residents, thus providing significant support for promoting common prosperity^[6]. Therefore, air quality is not only a key factor in addressing environmental and health crises, but also serves as a strategic pivot for driving high-quality economic development, promoting social equity, and achieving harmonious coexistence between humanity and nature.

Digital technology, as a core driver of economic transformation and upgrading, is profoundly reshaping modes of production, lifestyles, and the patterns of socioeconomic development. Over the past decade, rapid advancements in information and communication technologies have propelled the digital economy to become a primary engine and key driver of global economic growth. The outbreak of the COVID-19 pandemic has further accelerated the adoption and application of digital technologies, prompting a large-scale shift of work, daily life, and consumption activities online. The deep penetration of digital technologies has reconfigured production and lifestyles, and their role in promoting carbon emission reduction, driving high-quality development, and narrowing development gaps is increasingly prominent, making them a priority in national policy frameworks^{[7][8]}. In China, the growth momentum of the digital economy is particularly pronounced. In 2024, the value-added output of core digital economy industries reached 14.0891 trillion yuan, accounting for 10.5% of GDP—a 5.7% increase from the previous year—demonstrating its substantial contribution to the national economy^[9]. Policy initiatives have advanced in parallel. The Guidelines on Accelerating the Comprehensive Green Transition of Economic and Social Development, issued by the Central Committee of the Communist Party of China and the State Council in the same year, explicitly calls for strengthening green and low-carbon technological innovation and enhancing the quality, efficiency, and resource-energy utilization efficiency of economic and social development. This document charts the course for the deep integration of digital technology and the green transition^[10]. Academia has conducted extensive and in-depth research on the digital economy. Existing studies focus on its role in promoting post-pandemic sustainable development, encompassing its positive impacts on economic growth, resource use efficiency, and environmental protection^[11]. In recent years, the impact of the digital economy on air quality has gradually become a significant research topic. Empirical studies indicate that, overall, the development of the digital economy helps reduce air pollutant emissions and improve urban air quality^[12]. The core mechanism behind this improvement, however, is not the direct reduction of emissions. Instead, it is achieved by reshaping the economic structure and driving innovation. Specifically, the digital economy enhances the technical efficiency of pollution control by promoting green technological innovation. It promotes green technology innovation, thereby enhancing the technical efficiency of pollution control^[13]. Concurrently, it drives the advancement and optimization of industrial structure, fostering the growth of service and high-tech industries while reducing the share of traditional, heavily polluting, and high-emission industries, thus curbing emissions at source. Furthermore, through the synergistic effects of digital platforms, the digital economy optimizes resource allocation and reduces redundant energy consumption^[14]. Collectively, this research provides a crucial foundation for understanding the impact of the digital economy on air quality.

The above studies collectively indicate the positive role of digital economy development in improving air quality.

However, current studies on the impact of the digital economy on air quality still face several critical limitations that need to be addressed. Firstly, there are limitations in research scale. Most existing studies focus on macro-level analysis at the national or provincial level, with relatively few exploring more detailed scales such as the city level. This makes it difficult to reveal specific patterns such as local heterogeneity. Secondly, the existing literature pays insufficient attention to the spatial correlation between air quality and the digital economy itself, as well as the spatial heterogeneity arising from their interaction. This inadequacy hinders a deeper understanding of the complex relationship between the two.

2. Theoretical Analysis

2.1 Air Quality

Air quality denotes the concentration levels and integrated states of various air pollutants within a specific region. It is typically assessed based on environmental standards and health guidelines. In academic research and policy application, air quality is usually quantified by monitoring the concentrations of a set of key pollutants. Common core indicators include particulate matter and gaseous pollutants. Particulate matter, especially fine particulate matter (PM_{2.5}) and inhalable particulate matter (PM₁₀), is the primary culprit responsible for the formation of haze and poses a major threat to human respiratory health. It is also widely used as a proxy variable for air pollution in current research^[15]. Gaseous pollutants, such as sulfur dioxide (SO₂), nitrogen oxides (NO_x), carbon monoxide (CO), and ozone (O₃), predominantly originate from fossil fuel combustion and industrial processes. They serve as precursors to the formation of acid rain and photochemical smog. This paper focuses on “air quality” as a comprehensive environmental condition, which is characterized principally by the concentrations of the aforementioned pollutants.

2.2 Digital Economy

The digital economy refers to a series of economic activities that utilize digitalized knowledge and information as a key production factor, modern information networks as an important carrier, and the effective use of information and communication technologies to enhance economic efficiency and optimize the economic structure^[16]. Its core can be divided into two dimensions: digital industrialization and industrial digitalization. Digital industrialization refers to industries that provide technology, products, and services for digital transformation. These primarily include the electronic information manufacturing industry, the information and communication technology industry, the software and information technology service industry, and the internet industry. This constitutes the “hardcore” foundation of the digital economy. Industrial digitization, conversely, focuses on the increased output and enhanced efficiency derived from applying digital technologies in traditional industries, spanning a broad spectrum such as smart manufacturing, industrial internet, smart agriculture, e-commerce, digital finance, smart logistics, and telemedicine across the national economy. This dimension is central to the enabling role of the digital economy, driving the intelligent and efficient transformation of conventional sectors^[17]. Current academic and policy practices employ several mainstream methods to measure the development level of the digital economy. These mainly include the value-added approach, the index approach, the supply-demand perspective assessment approach, the international comparison and satellite account approach, and the regional disparity and structural analysis approach. The value-added approach, the most fundamental and widely adopted method, typically decomposes the digital economy into “digitalization of industry” and “industrial digitization,” separately estimating their contributions to economic growth and summing them to derive the digital economy’s added value^[18]. The index approach, on the other hand, constructs a multi-dimensional evaluation system that covers features such as digital industrialization, industrial digitalization, and the penetration, collaboration, and substitution of digital technologies. This method aims to comprehensively reflect both the breadth and depth of digital economy development^[19]. The supply-demand perspective assessment method offers a dual-path assessment from both supply and demand sides, addressing limitations of traditional unidirectional metrics and yielding more systematic and comprehensive results^[20]. The international comparison approach focuses on cross-national comparisons, often using developed economies like the US and Australia as references to assess China’s international competitiveness and relative position via indicators such as digital economy added value^[21]. The satellite account approach draws on the statistical frameworks of international organizations such as the Organisation for Economic Co-operation and Development (OECD). It treats the digital economy as a “satellite account” within the system of national economic accounting to specifically measure

its output scale and economic impact^[22]. The regional disparity approach primarily employs tools like the Gini coefficient and Dagum decomposition to analyze disparities in digital economy progression among China's eastern, central, and western regions and across different development tiers, revealing the spatial distribution of the digital divide^[23]. The structural analysis approach focuses on the internal industrial composition of the digital economy. By examining the proportional relationship between core industries and related supporting industries, as well as their dynamic evolution, this method captures the internal structure and development trends of the digital economy^[24].

2.3 Socioeconomic Factors

The relationship between economic development and environmental pollution is a central topic in environmental economics. The most classic theoretical framework is the Environmental Kuznets Curve (EKC) hypothesis. This hypothesis posits that in the early stages of economic development, environmental pollution intensifies as per capita income increases. However, once economic development reaches a certain level, the degree of environmental pollution gradually decreases with further income growth. This shift is driven by factors such as industrial structure optimization, technological progress, and increased public environmental awareness, presenting an inverted "U-shaped" relationship^[25]. In academia, per capita GDP is commonly employed as the core proxy variable for a region's economic development level^[26].

Urbanization, the concentration of population, capital, and land in cities, exerts a dual impact on air quality^[27]. Studies indicate increased urban population density leads surging energy consumption for daily living and transportation demand, thereby increasing pollutant emissions and degrading air quality^[28]. Concomitantly, rapid growth in vehicle ownership and traffic congestion substantially increases motor vehicle exhaust emissions^[29].

Industrialization is the most direct and primary driver of environmental pollution, particularly air pollution. Examining the energy consumption structure, relevant studies indicate that industrial production is the main consumer of energy, and China's energy structure has long been dominated by coal^[30]. Coal combustion emits significant amounts of SO₂, NO_x, and soot, key precursors and components of PM_{2.5}. Regarding industrial structure, studies note that secondary industries, represented by heavy chemicals, steel, and building materials, is a typical resource-intensive and pollution-intensive sector. A higher proportion of secondary industry in a region often correlates with a "crude" growth model and greater environmental pressure^[28]. Therefore, the level of industrialization, particularly its structure and quality, directly determines the intensity of pollution emissions in a region.

Government expenditure is a critical instrument for promoting environmental governance and reducing pollution and carbon. Research demonstrates that structural adjustments in fiscal spending can markedly enhance regional environmental governance capacity, with specific environmental protection expenditures significantly improving pollution control outcomes^[31]. In addition to direct environmental spending, expenditures on economic construction, science, education, culture, and health, as well as social security, can also improve the quality of regional development and indirectly support environmental governance^[32].

Scientific and technological innovation is the fundamental pathway for resolving environmental issues, primarily by fostering a "technique effect." Technological progress is core to improving energy efficiency, reducing energy consumption per unit of GDP, thereby mitigating pollutant emissions from energy use^[27]. Furthermore, innovation activities in a region can generate positive spillover effects on the green technology level of surrounding areas through channels such as talent mobility and technology diffusion^[33].

Opening-up influences the environmental quality of host countries through international trade and foreign direct investment (FDI). The environmental impact of FDI is heterogeneous^[34]. In technology-intensive and high value-added sectors, FDI typically brings positive environmental technology spillovers; in labor-intensive or resource-based industries, negative scale effects may still dominate^[35]. Meanwhile, influenced by international environmental pressures and domestic policy orientation, FDI is accelerating its shift away from traditional "high-pollution, high-energy-consumption" sectors toward "green energy, digital economy, and advanced manufacturing"^[36].

2.4 The Impact of the Digital Economy on Air Quality

The digital economy contributes to reducing air pollution^[37]. It takes data as a key production factor and relies on digital

infrastructure such as the Internet, 5G, and the Internet of Things. Through technology empowerment, it promotes the intelligent integration of industries and the development of virtualized services. By leveraging network effects for open sharing, it significantly enhances resource allocation efficiency and promotes low-carbon development^{[38][39]}. As a new type of production factor, data has the characteristics of low reproduction cost and a marginal cost of use approaching zero. This feature is often likened to “digital oil,” giving it an advantage in resource allocation efficiency that traditional factors cannot match^[40]. As economic activities extend from physical to virtual spaces, the operational virtuality of the digital economy reduces reliance on physical logistics and production, inherently enabling reductions in energy consumption and carbon emissions. Simultaneously, the digital economy relies heavily on high-speed, low-latency network infrastructure (such as 5G) and ubiquitously connected IoT systems. This enables real-time data collection and intelligent responses across all stages of production, circulation, and consumption, laying the technical foundation for refined resource management and precise environmental monitoring^[41]. Furthermore, digital technologies like artificial intelligence and big data are rapidly embedding into traditional manufacturing and services, promoting the proliferation of smart manufacturing and intelligent services. By optimizing processes, reducing waste, and enhancing output efficiency, they further drive low-carbon and intensive transformation of industrial chains^[42]. Moreover, the data openness, sharing, and collaborative mechanisms championed by the digital economy allow platform-based models to leverage network effects, effectively lowering market transaction costs and improving total factor productivity, thereby creating conditions for synergistic achievement of environmental and economic benefits.

3. Methodology

3.1 Data Sources

This study utilizes panel data for prefecture-level cities in China from 2020 to 2023. The primary data sources are the China City Statistical Yearbook, the “China City Digitalization Evolution Index (DEI) White Paper published by H3C, the Science Data Bank, and the statistical yearbooks of various provinces and prefecture-level cities.

3.2 Variable Selection

3.2.1 Dependent Variable

To comprehensively measure regional air pollution, the annual average of the Air Quality Index (AQI) for each prefecture-level city is selected as the dependent variable. The Air Quality Index is a composite metric for assessing air pollution severity, based on the concentrations of six key pollutants: PM_{2.5}, PM₁₀, SO₂, NO₂, O₃, and CO. The AQI is an important metric widely adopted internationally for the quantitative evaluation of ambient air quality. The data is derived from real-time monitoring at air quality stations across China between 2014 and 2024. The city’s daily AQI is first calculated from station-level daily averages, then aggregated to an annual average, and finally compiled into a dataset of annual AQI values for Chinese prefecture-level cities.

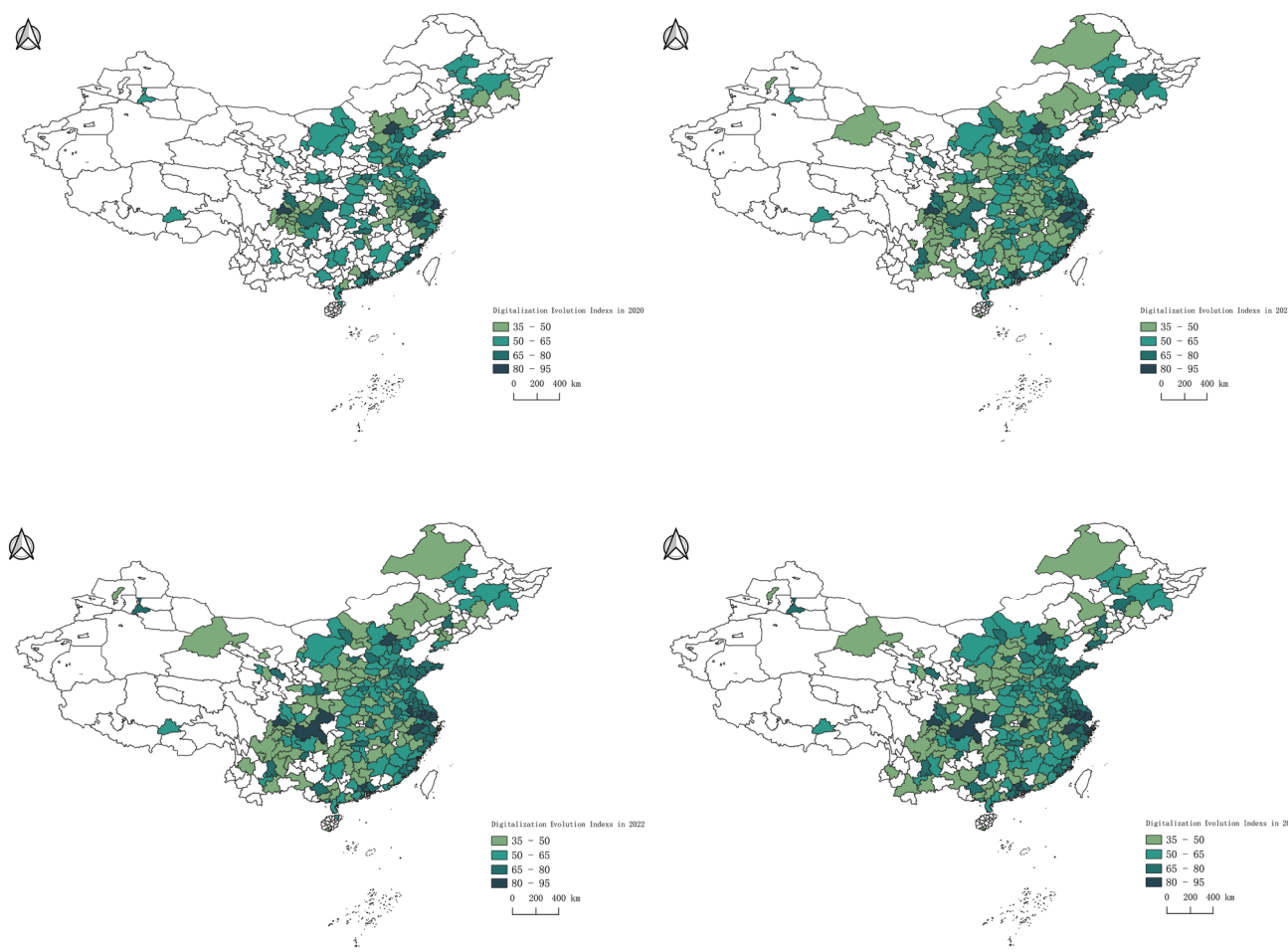
3.2.2 Core Explanatory Variable

To precisely characterize the comprehensive development level of the digital economy in each city, the core explanatory variable is the Urban Digitalization Evolution Index. Based on relevant literature, this study measures it using the Urban Digitalization Evolution Index (DEI) published in the “China City Digitalization Evolution Index (DEI) White Paper” released by H3C for the years 2020-2023. This index systematically evaluates the integrated development of the digital economy in major Chinese cities using a framework of four primary indicators—Data and IT Infrastructure, City Services, City Governance, and Industrial Integration—along with 12 secondary and 40 tertiary indicators. It is widely recognized and timely. The assigned weights are 20% for Data and IT Infrastructure, 35% for City Services, 20% for City Governance, and 25% for Industrial Integration. The white paper covered 147, 226, 231, and 240 prefecture-level cities in 2020-2023, respectively. This index is highly compatible with the prefecture-level city research scale of this paper and can be directly obtained and applied to empirical analysis.

Spatially, from 2020 to 2023, the development level of China’s urban digital economy consistently exhibited a pattern of “higher in the east and lower in the west, and higher in coastal areas than inland”. Specifically, high-value zones (index 80-95) are predominantly clustered in eastern coastal urban agglomerations such as Beijing-Tianjin-Hebei region, the Yangtze

River Delta, and the Pearl River Delta, forming core growth poles. Medium-high value zones (index 65–80) are widely distributed in the hinterlands of the eastern coast and surrounding provincial capitals in central and western China, indicating an initial diffusion of digital technology from core cities. In contrast, most cities in central, western, and northeastern regions remain in low-level development zones, forming contiguous “digital lowlands”. This pattern reveals significant spatial imbalance in China’s digital economy development, with digital dividends largely confined to a few developed regions, while many small and medium-sized cities and traditional industrial cities have not yet deeply integrated into the digitalization process.

Figure 1: Spatial Distribution of China’s Prefecture-Level City Digitalization Evolution Index, 2020-2023



3.2.3 Control Variables

To conduct a in-depth analysis of the digital economy’s impact on air quality, the following control variables are selected: (1) Economic development level, measured by per capita GDP of each city in the given year ; (2) Urbanization level, measured by the proportion of urban population in each city’s total population for that year; (3) Industrialization level, measured by the proportion of secondary industry in the regional GDP of each city; (4) Government intervention level, measured by the scale of local general public budget expenditure; (5) Technological innovation level, measured by the number of invention patents granted in each city for that year;(6) Openness level, measured by total import and export volume of each city for the year.

Table 1: Attributes of Control Variables

Variable Abbreviation	Variable Name	Measurement Method	Unit
EL	Economic Development Level	Annual per capita GDP	100 million yuan per person
UL	Urbanization Level	The proportion of the urban population in the total population each year	%

Variable Abbreviation	Variable Name	Measurement Method	Unit
LI	Industrialization Level	The proportion of the added value of the secondary industry in the regional GDP each year	%
GIL	Government Intervention Level	Scale of Local General Public Budget Expenditure by City	hundred million yuan
TIL	Technological Innovation Level	Number of authorized invention patents	piece
OL	Openness Level	Total import and export volume	hundred million yuan

3.3 Model Specification

3.3.1 Ordinary Least Squares

To preliminarily examine the overall directional impact and basic association between the digital economy and air quality, an Ordinary Least Squares (OLS) baseline model is first constructed. The model is specified as follows:

$$\ln AQI_{i,t} = \alpha + \beta_1 \ln DEI_{i,t} + \beta_2 \ln EL_{i,t} + \beta_3 \ln UL_{i,t} + \beta_4 \ln LI_{i,t} + \beta_5 \ln GIL_{i,t} + \beta_6 \ln TIL_{i,t} + \beta_7 \ln OL_{i,t} + \varepsilon_{i,t} \quad (1)$$

where $\ln AQI_{i,t}$ represents the air quality index of prefecture-level city i in year t , α is the intercept term, the core explanatory variable $\ln DEI_{i,t}$ is the logarithmic value of the Digitalization Evolution Index, and $\varepsilon_{i,t}$ is the random error term. Given that the macroeconomic environment and policies may fluctuate annually during the study period, the model essentially conducts separate regressions on four independent cross-sections from 2020 to 2023 to capture the distinct characteristics of each year.

3.3.2 Moran's I Analysis

To investigate whether a systematic spatial co-variation relationship exists between the digital economy and air quality, a bivariate Moran's I analysis is employed. Moran's I includes Global Moran's I and Local Moran's I. The Global Moran's I assesses whether spatial dependence is pervasive across the entire study area, while the Local Moran's I further identifies specific local spatial association patterns and clustering types for individual cities. Therefore, this study uses the bivariate Moran's I to reveal the spatial correlation between a region's digital economy development level and the air quality of its neighboring regions, thereby identifying their spatial dependency pattern.

The formula for the bivariate Global Moran's I is:

$$I_b = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(y_j - \bar{y})}{S^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad (2)$$

Here, $S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$, x_i denotes the Digitalization Evolution Index for city i , y_j denotes the Air Quality Index for city j , w_{ij} are the elements of the spatial weight matrix (constructed using the Queen contiguity criterion), and n is the number of spatial units. Moran's I ranges from $[-1, 1]$. If $I_b > 0$, it indicates a positive spatial correlation; if $I_b = 0$, there is no spatial correlation; and if $I_b < 0$, it signifies a negative spatial correlation.

Based on the global spatial autocorrelation analysis, it is possible to determine whether the digital economy and air quality exhibit a spatial clustering trend across the entire study area. Furthermore, by employing the local Moran's I, the heterogeneity in spatial association patterns among different cities can be revealed. This allows for the identification of locally significant cluster types, including four typical patterns: "High-High" clusters, "Low-Low" clusters, "High-Low" outliers, and "Low-High" outliers. This approach facilitates an in-depth analysis of the spatial synergistic or divergent characteristics between the development of the digital economy and air pollution. The formula for the Local Moran's I is:

$$I_i^l = \frac{(y_i - \bar{y})}{S^2 \sum_{j \neq i}^n w_{ij} (y_j - \bar{y})} \quad (3)$$

The local spatial autocorrelation index focuses on detecting small-scale spatial clustering or dispersion patterns. Unlike the Global Moran's I, it reveals spatial correlation for individual locations, analyzing clusters of high or low values within specific areas.

3.3.3 Geographically Weighted Regression

Geographically Weighted Regression (GWR) is a spatial regression analysis method that incorporates spatial variation into traditional global regression models. GWR uses environmental covariates to perform local linear regression on sample points

within a bandwidth range. It estimates a set of regression parameters at each location, thereby generating a series of local regression models. The formula for GWR is as follows:

$$\ln y_{GWR_i} = \alpha(\mu_i, v_i) + \beta_1(\mu_i, v_i) \ln DEI_i + \beta_2(\mu_i, v_i) \ln EL_i + \beta_3(\mu_i, v_i) \ln UL_i + \beta_4(\mu_i, v_i) \ln LI_i + \beta_5(\mu_i, v_i) \ln GIL_i + \beta_6(\mu_i, v_i) \ln TIL_i + \beta_7(\mu_i, v_i) \ln OL_i + \beta_8(\mu_i, v_i) \ln OL_i + \varepsilon_i \quad (4)$$

Here, $\ln y_{GWR_i} = 1$ is the predicted value of the model at location i ; $\alpha(\mu_i + v_i)$ is the intercept term; i denotes the i -th city; μ_i and v_i represent the longitude and latitude coordinates of the i -th city, respectively; and ε_i is the regression residual.

4. Analysis of Empirical Results

4.1 Analysis of Spatial Correlation

4.1.1 Bivariate Global Moran's I Analysis

This study employs the bivariate global Moran's I, with the Air Quality Index (AQI) as the benchmark variable, to examine its spatial association characteristics with both the Digitalization Evolution Index and various control variables. The test results are presented in the table below.

Regarding the core explanatory variable, the bivariate Moran's I between the Digitalization Evolution Index and Air Quality Index exhibits a trend of transitioning from negative to positive and strengthening annually over the sample period. In 2020, the Moran's I was -0.007, failing to pass the significance test. In 2021, the index turned positive to 0.011 and became significant at the 5% level. It continued to rise in subsequent years, reaching 0.025 in 2022 and further increasing to 0.028 in 2023. This evolutionary trend indicates that the spatial dependence between the digital economy and air quality is gradually forming and intensifying. Regions with a developed digital economy are increasingly co-located in space with surrounding areas experiencing higher pollution levels, suggesting a solidifying pattern of homogeneous agglomeration characterized by "high digitalization-high pollution."

For the control variables, most exhibit a significant positive spatial correlation with air quality. The Moran's I for economic development level has remained above 0.05 and statistically significant since 2021, indicating that areas with higher economic development tend to be surrounded by regions with similar air quality characteristics. Industrialization level and the level of government intervention both show a significant positive correlation with air quality across all years examined. This indicates that industrial agglomeration and the scale of government expenditure exhibit distinct spatial homogeneity, and they share a synergistic relationship with the regional distribution of air quality. The Moran's I for technological innovation level is significantly positive in 2020, 2021, and 2023, implying that areas with active innovation activities tend to cluster spatially and are associated with the air quality conditions in surrounding areas. The significance level for urbanization is relatively weaker but shows a significant positive correlation in 2022 and 2023. The Moran's I for the level of openness to foreign trade is negative in most years and significant in 2021 and 2023, reflecting that foreign trade-active regions may exhibit a pattern of heterogeneous spatial distribution.

These results collectively demonstrate the widespread existence of significant spatial autocorrelation in air quality and its influencing factors. Consequently, following the baseline regression, it is necessary to further introduce models capable of handling spatial dependence to more accurately identify the impact mechanism of the digital economy on air quality.

Table 2: LeBron's advantages in rebounds and assists are particularly obvious

Year	Variable	lnDEI	lnEL	lnUL	lnLI	lnGIL	lnTIL	lnOL
2020	lnAQI	-0.007	0.006	0.007	0.017**	0.014**	0.017*	-0.002
2021		0.011**	0.052*	0.007	0.010*	0.016**	0.087**	-0.008*
2022		0.025*	0.076*	0.018*	0.038**	0.032**	-0.004	0.001
2023		0.028**	0.062**	0.013*	0.061***	0.035***	0.100***	-0.008*

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

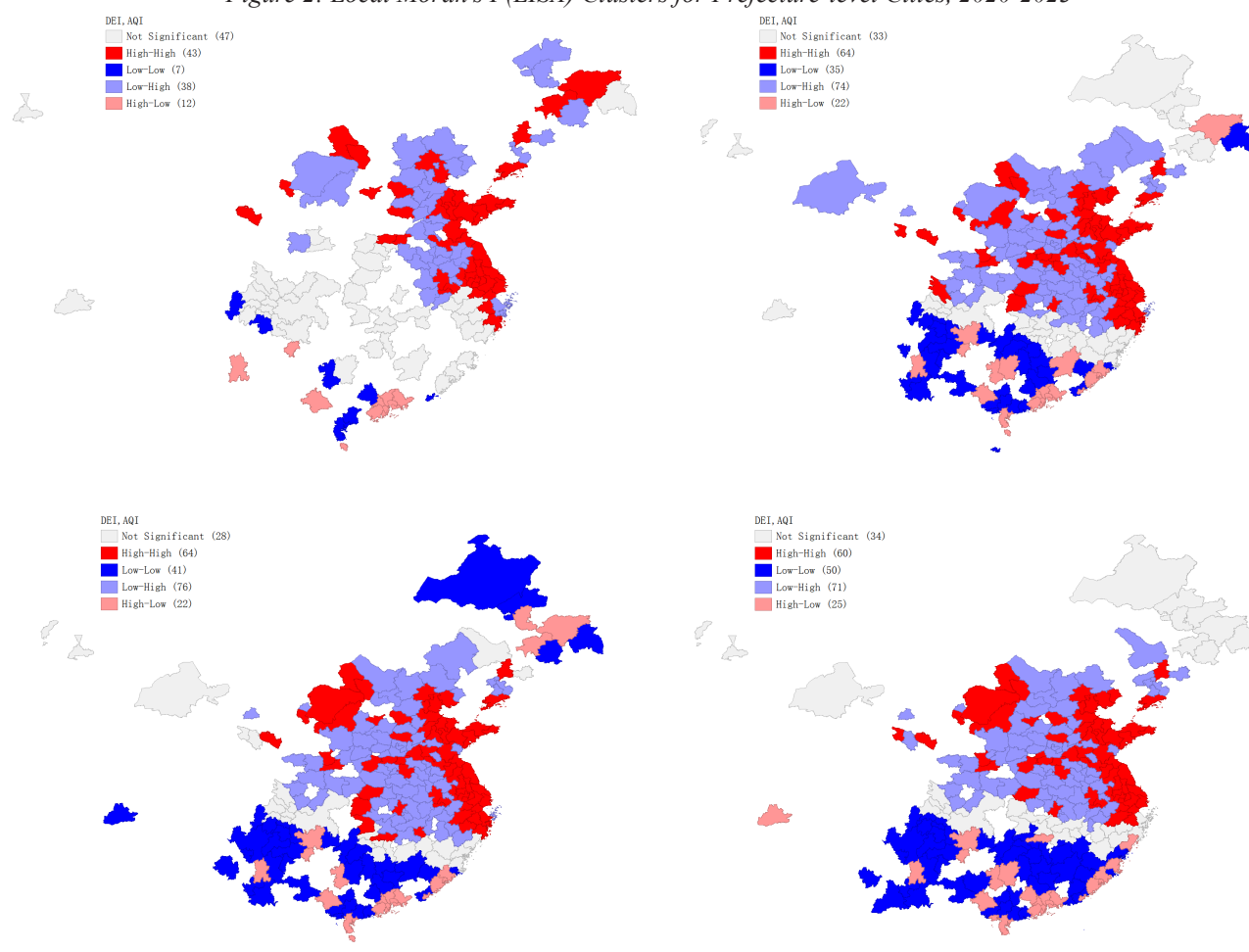
4.1.2 Local Moran's I Analysis

The analysis of Local Moran's I (LISA) can further reveal the specific spatial positions and cluster types of different cities

within the spatial association framework, providing a more detailed perspective for understanding the internal structure of the relationship between digital economic development and air quality. The figure presents the LISA cluster maps for the prefecture-level cities studied from 2020 to 2023.

In terms of the clustering pattern, the four types of spatial association modes show distinct quantitative differences and spatial differentiation. Among them, the “Low-High” type is the most numerous and stable in distribution, widely found around traditional industrial agglomerations and energy bases. The digital economy development level of these cities lags relatively behind, yet they bear the pressure of high air pollution from surrounding areas, demonstrating a distinct characteristic of “passive pollution bearing.” The “High-High” type ranks second in number. These cities are primarily located in regions where the digital economy is developing rapidly but has not yet fully decoupled from the heavy chemical industry, such as the Shandong Peninsula, the southern part of the Beijing-Tianjin-Hebei region, and some peripheral cities in the Yangtze River Delta. These areas reflect that the spatial overlap between the digital economy and pollution emissions remains relatively common, and the agglomeration of digital industries has not yet effectively translated into a driving force for air quality improvement. The “Low-Low” type is moderate, mainly distributed in regions with good ecological conditions and lower industrialization levels. Examples include some ecological function zones in southwest China and forestry cities in the northeast. In these areas, digital economic development is slow, but the air quality is inherently better, forming a state of relatively low-level equilibrium. The “High-Low” type is the smallest in number and shows no significant growth over time. This indicates that only a few cities have achieved a positive synergy between a leading digital economy and improved air quality in surrounding areas, with most of these being core cities in the Yangtze River Delta and the Pearl River Delta. The local spatial association analysis demonstrates that the relationship between the digital economy and air quality is not a simple linear one. It not only validates the robustness of the global conclusion but also reveals a multi-level and multi-type spatial dependence structure beneath the same global trend.

Figure 2: Local Moran's I (LISA) Clusters for Prefecture-level Cities, 2020-2023



4.2 Collinearity Test

Before conducting the regression analysis, it is necessary to diagnose whether there is a multicollinearity issue among the explanatory variables. When a high correlation exists between independent variables, it can lead to inflated standard errors of the coefficient estimates and unstable estimation results, thereby affecting the reliability of statistical inference. This paper employs the Variance Inflation Factor (VIF) to test for collinearity among the explanatory and control variables for each year. The table presents the VIF test results for the variables from 2020 to 2023. The findings indicate that all VIF values for all variables in each year are below 10, with most values concentrated between 2 and 6. Only the VIF for technological innovation level in 2023 is 7.38, slightly elevated compared to other years but still within an acceptable range. This indicates that there are no severe multicollinearity issues among the variables selected in this paper, allowing for subsequent baseline regression and Geographically Weighted Regression analysis.

Table 3: Multicollinearity Test

Variable	2020 VIF	2021 VIF	2022 VIF	2023 VIF
lnDEI	3.46	5.13	5.95	5.16
lnEL	3.22	2.72	2.86	3.44
lnUL	3.03	2.31	3.13	3.31
lnLI	1.38	1.41	1.16	1.46
lnGIL	3.28	2.99	3.13	3.31
lnTIL	4.79	4.95	5.95	7.38
lnOL	4.47	4.01	3.89	3.54

4.3 Comparison of Baseline and Geographically Weighted Regression Models

To clarify the correlation between the digital economy and local air quality across prefecture-level cities, this study conducts an empirical analysis based on the baseline regression model. Overall, the digital economy demonstrates an improving effect on air quality in each year, although its statistical significance fluctuates. The coefficients for 2021 and 2022 are significant at the 5% and 10% levels, respectively, while those for 2020 and 2023 are positive but fail to meet conventional significance thresholds. In terms of overall explanatory power, the R^2 values range from 0.09 to 0.18 annually, indicating that the OLS model explains a small portion of the variation in air quality. Thus, the baseline regression provides preliminary support for the digital economy's beneficial impact on air quality, but it has the following limitations: insufficient annual stability, limited explanatory capacity, and omission of spatial correlation controls.

However, the OLS model is inadequate for precise estimation. This section introduces the Geographically Weighted Regression (GWR) model and systematically compares its performance with that of the OLS model to statistically demonstrate the superiority of GWR in addressing spatial non-stationarity.

Comparing OLS and GWR model results for each year from 2020 to 2023 yields a clear and consistent conclusion: the Geographically Weighted Regression model comprehensively outperforms the Ordinary Least Squares model in characterizing the digital economy–air quality relationship.

Firstly, regarding goodness of fit, the explanatory power of the GWR model substantially exceeds that of OLS. Over the four-year span, the R^2 values of the OLS model range only from 0.094 to 0.179, implying that the traditional global regression can explain less than 18% of the variation in air quality at most, indicating a relatively poor model fit. In contrast, the R^2 values of the GWR model remain consistently high, between 0.837 and 0.858, accounting for over 80% of spatial variation in air quality and markedly enhancing real-data representation. This demonstrates a significantly enhanced ability to capture real-world data patterns. This substantial gap indicates that the impact of the digital economy on air quality exhibits pronounced spatial non-stationarity, and global regressions that ignore spatial heterogeneity risk losing a considerable amount of valuable information. Secondly, from a model adequacy perspective, the AICc values of the GWR model are substantially lower than those of the OLS model annually. The AICc is a commonly used criterion for balancing goodness of fit and model simplicity, where lower values indicate a better model specification. The data show that the AICc values of the GWR model are more than 140

points lower than those of the OLS model annually. This difference provides strong statistical evidence that the GWR model, which incorporates spatial heterogeneity, is a more scientifically sound specification. Thirdly, regarding residual magnitude, the fitted residuals of the GWR model are far smaller than those of the OLS model. For instance, in 2020, the residual of the OLS model is 131.154, while that of the GWR model drops to 23.895—an over 80% reduction. Similar patterns persist in other years, with GWR model residuals consistently stabilizing around 30, compared to OLS model residuals often exceeding 200. This significant reduction in residuals further confirms that the GWR model captures the underlying data characteristics far more effectively. Additionally, the bandwidth of the GWR model remains between 57 and 68 across all years, indicating that local estimation employs a moderate number of neighboring cities, which ensures an adequate capture of local characteristics while avoiding the problem of overfitting. Meanwhile, although the number of observations has gradually increased from 147 to 240 over the four-year period, with a continuously expanding sample coverage, the goodness of fit of the GWR model remains consistently stable. This demonstrates that the estimation results of the model possess strong robustness. Collectively, this analysis confirms that the impact of the digital economy on air quality is not globally homogeneous; rather, it exhibits significant spatial non-stationarity. Employing GWR model for local estimation enables a more authentic and nuanced revelation of the spatial differentiation in this relationship, thereby laying a solid methodological foundation for subsequent in-depth analysis of the heterogeneous effects across different regions.

Table 4: Comparison Between the OLS Model and the GWR Model, 2020–2023

Model	OLS				GWR			
Year	2020	2021	2022	2023	2020	2021	2022	2023
AICc	419.714	628.334	664.072	666.732	280.640	379.563	418.531	397.242
R ²	0.118	0.179	0.094	0.129	0.837	0.852	0.847	0.858
AIC	416.400	625.508	661.257	663.949	239.888	333.913	359.485	345.133
Residual	131.154	193.398	220.933	209.045	23.895	33.801	35.421	34.133
Observation	147	226	231	240	147	226	231	240
Bandwidth					57	68	58	65

4.4 Spatial Differentiation Analysis of the Digital Economy's Influence Coefficients

This section further visualizes the distribution of the influence coefficients of the digital economy on air quality, as estimated by the GWR model from 2020 to 2023, across different prefecture-level cities. This allows for the identification of the spatially differentiated performance of the digital economy's environmental effects.

To intuitively present the spatial distribution differences in the influence coefficients of the digital economy on air quality, this paper adopts the method of “mean \pm multiplier \times standard deviation” to classify the local GWR estimation coefficients of each prefecture-level city. Specifically, using the mean of the influence coefficients of all sample cities as the benchmark, and taking 0.5 and 1.5 times the standard deviation as the critical thresholds, all cities are divided into five categories: regions below the mean minus 1.5 times the standard deviation are defined as significant inhibition zones, indicating that the digital economy has a significant negative impact on air quality; regions between the mean minus 1.5 times and the mean minus 0.5 times the standard deviation are defined as low-value zones, indicating that the digital economy has a weak positive or a slight negative effect on air quality improvement; regions between the mean minus 0.5 times and the mean plus 0.5 times the standard deviation are defined as medium-value zones, indicating that the impact is at an average level; regions between the mean plus 0.5 times and the mean plus 1.5 times the standard deviation are defined as medium-high-value zones, indicating that the digital economy has a relatively obvious effect on improving air quality; regions above the mean plus 1.5 times the standard deviation are defined as significant promotion zones, indicating that the digital economy has the most significant effect on improving air quality.

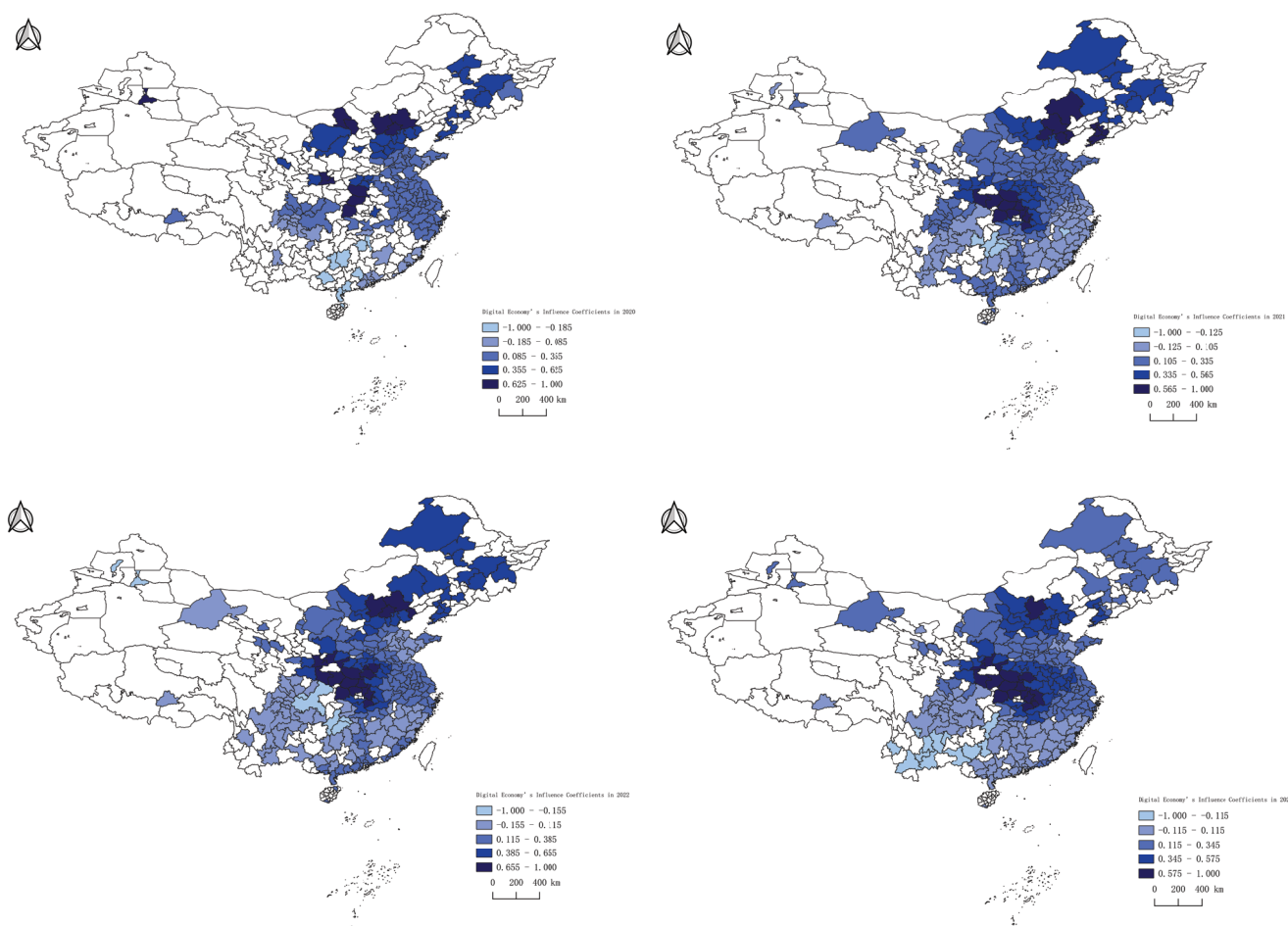
From an overall spatial perspective, the influence coefficients of the digital economy on air quality exhibit a distinct east-west

divergence, a pattern that remains relatively stable throughout the sample period. The figure presents the spatial distribution of these influence coefficients from 2020 to 2023, mapped according to the classification criteria outlined above. Specifically, the medium-high-value zones and significant promotion zones are mainly concentrated in the eastern coastal areas, particularly in the core cities of the Yangtze River Delta, the Pearl River Delta, and the Beijing-Tianjin-Hebei region. These areas have a leading level of digital economy development, and the improving effect of the digital economy on air quality is most prominent here. In contrast, the low-value zones and even significant inhibition zones are widely distributed across the central and western regions, covering most prefecture-level cities in Northeast China, the western part of North China, the Northwest, and the Southwest. In these regions, the influence coefficients are negative or only weakly positive. This indicates that the environmental dividends of digital technology have not yet been effectively realized, and in some areas, there is even a coexistence of digital expansion and pollution accumulation.

Examining temporal evolution trends from 2020 to 2023, the environmental effects of the digital economy are undergoing a gradient diffusion from core cities to peripheral regions, although the pace of this diffusion remains relatively slow. The scope of the significant promotion zones is expanding gradually. Some secondary coastal cities, such as Qingdao and Yantai on the Shandong Peninsula, and Quanzhou and Zhangzhou in Fujian Province, have progressively advanced from medium-value zones to medium-high-value or significant promotion zones. Concurrently, while the extent of significant inhibition zones has contracted somewhat, vast areas in central and western China remain covered by low-value zones. This is particularly evident in the northwestern region and the old industrial bases in the northeast, where the degree of improvement has been relatively limited.

The spatial differentiation pattern described above reinforces the core judgment made earlier: the impact of the digital economy on air quality is not globally homogeneous. This finding also provides clear geographical targeting for the policy recommendations that follow.

Figure 3: Spatial Distribution of the Digital Economy's Influence Coefficients Across Chinese Prefecture-Level Cities, 2020–2023



5. Conclusions and Policy Recommendations

5.1 Conclusions

First, there is a significant spatial dependence between the digital economy and air quality, and this correlation pattern has undergone a directional shift and continued strengthening during the sample period. The global Moran's I shows that the spatial correlation between the two gradually transitioned from a weak negative correlation in 2020 to a positive spatial agglomeration after 2021, exhibiting a co-agglomeration characteristic of "high digital economy - high air pollution." Meanwhile, most control variables, such as the level of economic development, industrialization, government intervention, and technological innovation, also show significant positive spatial correlations. This indicates that these factors have distinct spatial homogeneity in their regional distribution and exhibit a synergistic relationship with the regional pattern of air quality. Second, the local spatial association analysis reveals a more complex internal structure. The vast majority of cities in China remain in a state of spatial overlap between the digital economy and environmental pollution, or are passively bearing pollution pressure, indicating that achieving coordinated development between the two still has a long way to go. Among the four types of agglomeration patterns, "Low Digital Economy - High Air Pollution" cities are the most numerous. Widely distributed around traditional industrial agglomerations and energy bases, these cities exhibit a distinct characteristic of "passively bearing pollution," where their own digital economy development lags while they suffer from high pollution pressure originating from surrounding areas. The "High Digital Economy - High Air Pollution" type ranks second. These cities are mainly located in regions where the digital economy is developing rapidly but has not yet decoupled from the heavy chemical industry. "Low Digital Economy - Low Air Pollution" cities are concentrated in areas with a good ecological environment and a low degree of industrialization. In contrast, "High Digital Economy - Low Air Pollution" cities, representing a positive synergy, are extremely rare and are only sparsely distributed in core cities such as the Yangtze River Delta and the Pearl River Delta.

Thirdly, model superiority of the Geographically Weighted Regression (GWR) is confirmed. The explanatory power of the traditional global OLS model is limited. In contrast, the GWR model demonstrates significant advantages, with a goodness-of-fit (R^2) consistently between 0.837 and 0.858 across the four years, explaining over 80% of air quality variation, compared to a maximum R^2 below 0.18 for OLS. Furthermore, the GWR model's corrected Akaike Information Criterion (AICc) values were, on average, more than 140 points lower, and residuals decreased from above 200 to around 30. This underscores that incorporating spatial heterogeneity substantially improves model accuracy and better reflects the true data structure, confirming significant spatial non-stationarity in the impact of the digital economy on air quality and the potential for global models to obscure localized realities.

Fourthly, the influence coefficients of the digital economy estimated by the GWR model reveal a clear pattern of spatial differentiation. In the eastern coastal areas, particularly the core cities of the Yangtze River Delta, the Pearl River Delta, and the Beijing-Tianjin-Hebei region, the digital economy has the most significant effect on improving air quality, forming "significant promotion zones." In contrast, vast areas in central and western China are covered by "low-value zones" or even "significant inhibition zones." Although the scope of the significant promotion zones expanded somewhat between 2020 and 2023, the pace of diffusion has been relatively slow. This pattern indicates that in the developed eastern regions, the digital economy has begun to release green dividends, while in most central and western regions, it remains in a phase characterized by scale expansion coupled with energy consumption.

5.2 Policy Recommendations

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

First, implement differentiated regional governance strategies based on the type of spatial association. For the most numerous "Low Digital Economy - High Air Pollution" cities, efforts should focus on addressing shortcomings in digital infrastructure and promoting the low-cost adoption of green digital technologies to alleviate their predicament of passively bearing pollution. For "High Digital Economy - High Air Pollution" regions, it is necessary to accelerate the substantive decoupling of the digital economy from polluting industries. This can be achieved by tightening environmental access thresholds, implementing carbon performance evaluations, and employing other means to compel the green transformation of traditional

industries. For “High Digital Economy - Low Air Pollution” core cities, their experiences in integrating digital governance with environmental regulation should be promptly summarized to form replicable institutional achievements. For “Low Digital Economy - Low Air Pollution” ecological areas, ecological priorities should be upheld, with a moderate development of environmentally friendly digital industries while avoiding blind expansion.

Second, acknowledge spatial spillover effects and establish a framework for cross-regional collaborative governance. The “High-High” co-agglomeration of the digital economy and air pollution is not an isolated local phenomenon; rather, it is the result of the interplay between interregional factor flows and the cross-boundary transmission of pollution. It is recommended to establish “Digital-Environment” collaborative governance zones, using city clusters and metropolitan areas as the basic units. This would promote the sharing of emissions data, joint judgment of pollution sources, and shared responsibility for emission reduction. Simultaneously, guide the strategic distribution of digital infrastructure, such as data centers and computing hubs, to western regions rich in clean energy. This would help alleviate the spatial binding of “High Digital Economy” and “High Air Pollution” at the source.

Third, respect spatial non-stationarity to facilitate the precise adaptation of local policies. The estimation results of the GWR model indicate that the same set of digital economy policies can produce vastly different environmental outcomes in different cities. Therefore, the top-level national design should leave sufficient room for local adaptation, avoiding the assessment of all regions using uniform indicators. Local governments should be encouraged to conduct “Environmental Impact Pre-assessments for the Digital Economy.” Before the implementation of major digital industry projects, spatial econometric methods should be used to simulate their potential impact on local and neighboring air quality, providing scientific support for project site selection and scale regulation.

Fourth, align with the gradient diffusion pattern of the digital economy’s environmental effects to promote the orderly transfer of green dividends from the eastern to the western regions. Support for digital technology transfer and green innovation in the central and western regions should be strengthened. This can be achieved by establishing cross-regional digital transformation guidance funds and creating “Digital Pairing” assistance mechanisms between the eastern and western regions, thereby facilitating the adoption and adaptation of mature smart environmental protection and intelligent manufacturing solutions from the east in the central and western regions. At the same time, as the central and western regions undertake industrial transfers from the east, advanced air quality monitoring and treatment technologies should be introduced in tandem. This will prevent “digital divides” from becoming new “pollution sinks” and truly realize the cross-regional sharing of digital dividends and air quality benefits.

Funding

No

Conflict of Interests

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

Reference

- [1] World Health Organization.(2006). Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide.
- [2] Ministry of Ecology and Environment of the People's Republic of China. (2022). Report on the State of China's Ecology and Environment.
- [3] State Council of the People's Republic of China.(2023) .Notice on Issuing the “Action Plan for Continuous Air Quality Improvement” .
- [4] The Central People's Government of the People's Republic of China. (2025). Proposal of the Central Committee of the Communist Party of China on Formulating the 15th Five-Year Plan for National Economic and Social Development.
- [5] Li, Y., & Han, X. (2025). Coupling Coordination Effect between Regional Economy and Air Quality: An Empirical Test Based on Panel Data from the Beijing-Tianjin-Hebei Urban Agglomeration. *Journal of China University of Petroleum*, 41(06), 1-10. <https://doi.org/10.13216/j.cnki.upcjess.2025.06.0001>.
- [6] Zhang F, Zhou L, & Li H. (2025). Air Quality Monitoring and Urban-Rural Health Disparities. *Economics* , 25(05),

- 1219-1236. <https://doi.org/10.13821/j.cnki.ceq.2025.05.07>.
- [7] Qin, D., & Puttanapong, N. (2025). Estimating County-level Digital Economy Development in China using Satellite Data: Insights from Explainable Spatial Machine Learning. *Geocarto International*, 40(1), 2596971. <https://doi.org/10.1080/10106049.2025.2596971>.
- [8] Zhang J, Shi D, Dong Z, et al. (2024). The Carbon Emission Effects of Digitalization: A Literature Review and Research Prospects. *Journal of Economic Management*, 3(03), 31-56. <https://doi.org/10.20180/j.qjem.2024.03.02>.
- [9] Central People's Government of the People's Republic of China. (2025). Share of Core Digital Economy Value Added in GDP Reached 10.5% in 2024.
- [10] Central People's Government of the People's Republic of China. (2024). 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.
- [11] Le D. (2023). A Study on the Impact of the Digital Economy on Environmental Quality (Doctoral dissertation). Zhongnan University of Economics and Law. <https://doi.org/10.27660/d.cnki.gzczu.2023.000160>.
- [12] Zhang J. (2023). Research on the Spatial Effects and Regional Heterogeneity of the Digital Economy's Impact on Air Pollution (Master's thesis). Inner Mongolia University of Finance and Economics. <https://doi.org/10.27797/d.cnki.gnmgc.2023.000534>.
- [13] Sun Q. (2024). Analysis of the Impact of the Digital Economy on Nitrogen Oxide Emission Pollution in Air: An Empirical Study Based on Interprovincial Panel Data from 2013 to 2020. *Agricultural Outlook*, 20(12), 47-55.
- [14] Huang Q., & Cai X. (2024). Digital Economy Development and Air Quality Improvement: An Analysis Based on the Innovation-Driven Perspective. *Enterprise Economy*, 43(07), 91-101. <https://doi.org/10.13529/j.cnki.enterprise.economy.2024.07.009>.
- [15] Xu, S.; Yang, C.; Huang, Z.; Failler, P. (2022). Interaction between Digital Economy and Environmental Pollution: New Evidence from a Spatial Perspective. *International Journal of Environmental Research and Public Health*, 19, 5074. <https://doi.org/10.3390/ijerph19095074>.
- [16] Han, X. (2024). An Empirical Study on The Impact of Digital Economy on Industrial Pollution Emissions. *Frontiers in Business, Economics and Management*.
- [17] Yan S., & Lai X. (2025). Theoretical Mechanism and Empirical Test of the Digital Economy Promoting the Integrated Development of “Two Industries”: Based on Empirical Evidence from 50 Countries Worldwide. *Journal of Jimei University*, 28(01), 75-87.
- [18] Qin, D., & Puttanapong, N. (2026). Evaluating LJ1-01 Nighttime Light Data for Economic Analysis at China's County Level. *Remote Sensing Letters*, 17(3), 324-335. <https://doi.org/10.1080/2150704X.2026.2629349>.
- [19] Fan H., & Wu T. (2020). Measurement of China's Digitalization Level and Construction of an Indicator System. *Journal of Capital University of Economics and Business*, 22(04), 3-12. <https://doi.org/10.13504/j.cnki.issn1008-2700.2020.04.001>.
- [20] Gao Y., & Xu Z. (2021). Construction and Measurement of China's Digital Economy Evaluation System from Supply and Demand Perspectives. *Commercial Economy Research*, (16), 180-183.
- [21] Xu X., & Zhang M. (2020). Research on the Measurement of China's Digital Economy Scale: Based on an International Comparative Perspective. *China Industrial Economics*, (05), 23-41. <https://doi.org/10.19581/j.cnki.ciejournal.2020.05.013>.
- [22] Xu L., & Wang K. (2024). A Review of Research on the Connotation and Measurement Indicator System of the Digital Economy. *Statistics and Decision*, 40(12), 5-11. <https://doi.org/10.13546/j.cnki.tjyj.2024.12.001>.
- [23] Cheng G., & Li Y. (2022). Measurement of Digital Economy Development Level and Research on Regional Differences Based on Techno-Economic Paradigm. *Journal of Industrial Technological Economics*, 41(06), 35-43.
- [24] Guan H., Xu X.n, Zhang M., & Yu X. (2020). Research on Statistical Classification Issues of China's Digital Economy Industries. *Statistical Research*, 37(12), 3-16. <https://doi.org/10.19343/j.cnki.11-1302/c.2020.12.001>.
- [25] Ai X., Chen L., & Zhu L. (2017). Research on the Relationship Between Air Pollution Emissions and Economic Growth: Based on a Spatial Econometric Model of Chinese Provincial Panel Data. *East China Economic Management*, 31(03), 69-76.

- [26] Li Q., Song J., Zhang Ji., Yu W. & Hu H. (2013). Study on the Evolutionary Patterns of the Impact of Urbanization on Ambient Air Quality in China. *Acta Scientiae Circumstantiae*, 33(09), 2402-2411. <https://doi.org/10.13671/j.hjkxb.2013.09.023>.
- [27] Jiang L., Zhou H., & Bai L. (2019). The Relationship Between Urbanization Development and Urban Air Quality in China Based on a Spatial Econometric Model. *Tropical Geography*, 39(03), 461-471. <https://doi.org/10.13284/j.cnki.rddl.003142>.
- [28] Ye S., Wang P., Huang Y., Zhe Y., & Ding M. (2023). Impact of Urban Spatial Morphology on the Spatial Heterogeneity Characteristics of PM_{2.5} and O₃ Pollution in the Yangtze River Delta Urban Agglomeration. *Ecology and Environmental Sciences*, 32(10), 1771-1784. <https://doi.org/10.16258/j.cnki.1674-5906.2023.10.006>.
- [29] Wang F., Zheng J., & Wang Z. (2019). Analysis of the Impact of Urbanization Level and Industrialization Level on Air Quality: A Spatiotemporal Model Based on Data from 16 Cities in Hubei Province from 2005 to 2017. *Resources and Environment in the Yangtze Basin*, 28(06), 1411-1421.
- [30] Wei G., Yang Y., Li R., Liu Y., He B. (2024). Digital Economy Exhibits Varying Degrees of Mitigation of Air Pollution in China: Total Cities-Economic Subdivisions-Urban Agglomerations. *iScience*, 27(6), 110091. <https://doi.org/10.1016/j.isci.2024.110091>.
- [31] Qin, D., & Thepmongkol, A. (2025). The Economic Impact of Financial Integration: A Case of Chengdu-Chongqing Economic Circle. *Thailand and The World Economy*, 43(3), 142-161.
<https://so05.tci-thaijo.org/index.php/TER/article/view/267034>.
- [32] Ma, R., & Bo, N. (2023). Research on the Influence Mechanism and Effect of Fiscal Expenditure Structure on the High-Quality Development of the Yangtze River Delta Urban Agglomeration. *Finance and Trade Research*, 34(04), 69-82+110. <https://doi.org/10.19337/j.cnki.34-1093/f.2023.04.007>
- [33] Jiang, L., & Ji, M. (2011). Urbanization, Regional Innovation Clusters, and Spatial Knowledge Spillovers: An Empirical Study Based on a Spatial Econometric Model. *Soft Science*, 25(12), 86-90.
- [34] Zhang, Q., Zhang, R., & Zhang, Y. (2019). The Impact of Foreign Direct Investment on Environmental Quality under Environmental Regulation: A Comparative Study Based on Different Industry Groups. *Commercial Research*, (05), 61-68. <https://doi.org/10.13902/j.cnki.syyj.2019.05.008>
- [35] Jiang, L., Zhou, H., & Bai, L. (2018). Spatial Heterogeneity Analysis of the Impact of Foreign Direct Investment on Air Pollution: A Case Study of the Air Quality Index (AQI) in 150 Chinese Cities. *Scientia Geographica Sinica*, 38(03), 351-360. <https://doi.org/10.13249/j.cnki.sgs.2018.03.004>
- [36] Sheng, B., & Lyu, Y. (2012). The Impact of Foreign Direct Investment on China's Environment: An Empirical Study Using Industrial Panel Data. *Social Sciences in China*, (05), 54-75+205-206.
- [37] Guo, B., Wang, Y., & Zhang, H. (2022). Does the Development of the Digital Economy Improve Urban Air Quality? A Quasi-Natural Experiment Based on National Big Data Comprehensive Pilot Zones. *Journal of Guangdong University of Finance & Economics*, 37(01), 58-74. <https://doi.org/10.20209/j.gxcb.441711.2022.01.005>
- [38] Wang, J. (2024). Research on the Impact of Digital Economy Development on Carbon Emissions in Chinese Prefecture-Level Cities (Doctoral dissertation). Shandong Normal University. <https://doi.org/10.27280/d.cnki.gsdsu.2024.000023>
- [39] Lian, Y. (2023). Research on the Digital Economy Driving the Transformation Between Old and New Growth Drivers of China's Economic Development in the New Era (Master's thesis). Northwest University.
- [40] Wei, Z. (2021). Research on the Connotation and Characteristics of the Digital Economy. *Journal of Beijing Vocational College of Economic Management*, 36(02), 3-10.
- [41] Gao, F. (2025). Cultivating New Drivers and Advantages for Chinese Modernization through Digital Economy Development: A Review of Professor Ren Baoping's Monograph Digital Economy Empowers Chinese Modernization. *Knowledge and Action*, (05), 92-96. <https://doi.org/10.19309/j.cnki.zyx.2025.05.006>
- [42] Huang, Q., Yu, Y., & Zhang, S. (2019). Internet Development and Manufacturing Productivity Growth: Internal Mechanisms and Chinese Experience. *China Industrial Economics*, (08), 5-23. <https://doi.org/10.19581/j.cnki.ciejournal.2019.08.001>

An Integrated Financial–Operational Framework for Capital Equipment Decisions: Advancing Manufacturing Investment Theory and Practice

Hamza Saad*

School of Industrial Sciences and Technology, University of Central Missouri, MO, USA

*Corresponding author: Hamza Saad, saad@ucmo.edu

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 study combines both operational and financial perspectives of capital investment decisions in the manufacturing field, comparing a manufacturer's choice of acquiring a new laser cutting machine against continuing to use their current machinery. The study was conducted over a 6-month period in 2025, where two types of finance were analyzed: a single payment of \$700,000 as opposed to two payments of \$500,000 at the end of Year 1 and Year 2. A discounted cash flow analysis was completed for both types of finance with discount rates (5% - 12%) to measure the financial efficiency of both finance types, while the operational analysis analyzed all costs associated with labor, energy, maintenance and capital costs in order to determine the cost per unit produced. Upfront payment method was the most financially efficient option. At an 8% discount rate, the combined present value of the two installments (\$891,632) was \$191,632 greater than the upfront cost with 27.4%. The operational analysis concluded that the new laser cutting machine had a significantly greater uptime (95%) compared to the legacy machine's (60%) uptime and had much lower scrap (2%) compared to the legacy machine's (5%). However, the total cost per unit produced by the new laser cutting machine (\$13.16) was almost double that of the legacy machine (\$6.74). The developed integrated framework is transferable across all industries and therefore can be used by manufacturers of all types and sizes to align strategic financial feasibility with operational efficiency when making investment decisions.

Keywords: Capital Equipment Investment; Present Value of Cost (PVC); Operational Efficiency; Cost-Per-Unit Analysis; Manufacturing Decision-Making

Published: Mar 22, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1176>

1. Introduction

Investments in decisions for capital in manufacturing environments are critical drivers of a business's competitiveness, sustainability, and operational efficiency long term (Rahman, 2025). The technologies and stabilization strategies selected to be deployed impact not only production performance today, but also durability for tomorrow (Amjad et al., 2010). When making any technology investment decisions, a commitment to invest in capital technology such as a laser cutting machine is especially significant because of the impact it will have on productivity, standardization, and rapid change within the production systems (Rüßmann et al., 2015).

These capital investment decisions of adopting advanced technology are also limited by capital constraints (Lanteri & Rampini, 2025; Hobdari et al., 2009). Companies must consider various ways to finance the acquisition, such as paying for

infrastructure upfront in the form of a capital expenditure vs payments made in installments, which all have varying levels of cash flow implications, total cost of ownership and opportunity cost implications (Zeidan & Shapir, 2017).

Literature has offered considerable insight into the adoption of technology and capital financing strategies, yet there is limited clarity about incorporating financial evaluation and operational efficiency in real world decision making (Driouchi & Bennett, 2012). Studies examine capital budgeting procedures or operational performance, but not closely enough how these two evaluations could be combined to achieve a better evaluation.

There is an inherent research gap in which firms might select financially inefficient strategies for only marginal operational benefits. Numerous reports now demonstrate that small- and medium-sized manufacturers tend to prioritize short-term liquidity preservation over long-term cost effectiveness, which leads to higher financing costs and lower profitability (Driouchi & Bennett, 2012; Johnson & Scott, 2019). These examples often occur in local contexts where firms are financially constrained or have less access to affordable financing options.

This article studies a manufacturing company that assessed purchasing a laser cutting machine in a two-payment loan of \$500,000 (\$1,000,000 total) over 2 years, or for \$700,000 cash payment. The laser machine cuts faster but has remained idle during long cutting times that limit through-put efficiency, as it operates with more uptime than the company's existing cutting machine (Rahman, 2025; Lanteri & Rampini, 2025).

Furthermore, the study will assess the company's decision through financial modelling and operational analysis. It offers a knowledge contribution to managing technology in the larger setting of manufacturing technology management. It contributes to our understanding of how firms can align technology acquisition decisions with long-term competitive position. By synthesizing financial and operational measures in a single decision lens, the study goes beyond capital budgeting in the traditional sense, to a technology management lens. The dual lens approach illustrates the decisions to invest in technologically advanced machinery directly correlate with manufacturing resiliency, operational agility, and strategic competitiveness. Two questions will be guided by:

RQ1. From a financial perspective, which payment plans are more cost-effective to the operation considering whatever assumptions we take regarding discount rates?

RQ2. From an operational perspective, is there sufficient productivity advantage to the laser cutting machine given the additional financial cost brought on by the installment payment financing?

This research is firmly designed to advance the field's theory of investment appraisal by explicitly linking financial feasibility with operational capability. The paper amalgamates discounted-cash-flow reasoning with cost-per-good-unit assessments, thereby creating a connection between capital-finance decision theory and manufacturing-systems performance theory. In this comprising work, it situates itself firmly within the field of manufacturing strategy and Industry 4.0 literature, evidencing how the financial structure directly affects process capability and, subsequently competitiveness in the long term. These topics are of major concern for the Journal of Manufacturing Technology Management. Two hypotheses were developed to answer these questions. The first is whether financing structure has a significant effect on present value of the investment and the second whether a larger capacity technology will have a lower cost-per-good unit. The hypotheses guide the analysis and provide a context for interpreting findings.

H1 (Financing and Present Value)

Installment financing for new manufacturing equipment will have a higher present value of costs than an up-front, lump-sum purchase.

H2 (Technology and Cost-Per-Unit):

The new laser-based technology will not have a cost-per-good unit advantage over the incumbent technology if financing premiums and operating realities are accounted for despite nominal-capacity advantages.

The research focuses on manufacturing systems and industries instead of specific problem-solving contexts. The context of the study is used to demonstrate the concept of a generalizable decision-making framework that links the types of financing used to the type of manufacturing performance achieved. By clearly linking the financial form to the annualized capital cost and to the cost per good unit produced, this research provides a useful model for understanding the relationship between the

way investments are made and the ongoing productivity and cost efficiency of all capital-intensive manufacturing operations.

1.1 Study contribution

This work advances the theory of manufacturing decision-making by explicitly developing and validating an integrated framework for discounted cash-flow analysis incorporating cost-per-good-unit metrics. In contrast to previous manufacturing decision making work that distinguishes between financial and operational evaluations, the framework links capital financing structures with associated manufacturing performance outcomes including throughput, scrap, and uptime. By applying the framework in an SME laser cutting context, we demonstrate model application and show funding premiums can negate operational upside. Importantly, the framework is transferrable across industries affording a set of generalizable decision rules for manufacturers to estimate capital funding appendages associated with any capital investments. In so doing, the study further extends manufacturing theory associated with a total cost of ownership and investment appraisal by integrating operational realities into financial evaluations.

2. Literature review

The contribution of the paper is solid because it brings together financial discounted cash flow analysis with operational cost-per-unit analysis to fill an existing gap in the research about investment decisions in manufacturing. Many sources prior to this paper treat financial evaluations separately from operational evaluations - for example, while (Ellram, 1993) did introduce a Total Cost of Ownership (TCO) framework, which is fundamentally about the capacity to jointly take capital costs and operational costs into account, it certainly is a more rigid view than what is represented in this paper and conducted further which this paper has methodological contributions to investigate further. Similarly, as (Mellichamp, 2013) described, there are ways to notionally extend discounted cash flow methods to address uncertainty in investment evaluations, which relates to our focus on PVC and PVC sensitivity analysis.

The results presented in this study also connect to the real-options literature on technology adoption. (Driouchi & Bennett, 2012) described how managers exaggerate perceived productivity benefits and under-represent financing risks (often both), which leads to incorrect technology adoption decisions. In our paper, the evidence to show that installment financing does not provide any productivity-adjusted value over upfront payment supports this argument. Furthermore, (Hobdari et al., 2009) describe how firms in transition economies are often pushed toward high-cost financing structures as a result of financial structures which erode profitability - this finds an analogous structure with the penalty for installments in this article.

From an operations perspective, we believe our inclination in relation to value found with increased uptime of the laser cutter (95%) and reduced scrap generated (2%) in its full (or BPM) implementation remains at odds with capital intensity to be in keeping with (Baumers et al., 2016), which found that advanced manufacturing technology typically produced unacceptable economies of scale despite technical improvements. Likewise, (Shaw et al., 2004) described a performance model of lasers and showed that, in terms of potential productivity gains, possible productivity losses from cycle inefficiencies more often outweighed productivity gains, which would support our observations within operations.

However, what may seem limitations here can instead be framed as informing the strengths of this work. To begin with, the case-based design has offered rigorous analytical depth and replicate-ability but also offers a solid base for industry comparative validation, as proposed by (Burlea-Schiopoiu & Mihai, 2019) as part of their integrated SME sustainability framework. Second, the two-year fiscal horizon was indeed narrow, and intentionally so; a two-year time horizon highlights short-term liquidity considerations that are particularly relevant to SME survival. Ultimately, (Markus & Rideg, 2021) determined just how integral cash flow analysis is to competitive behavior. Third, leaving the long-term learning effects of strategic pathways open is not a gap, it is a deliberate opening, and as (Gherghina et al., 2020) illustrate, SMEs in contexts such as Horizon 2020 are indeed more flexible and adaptive. In addition, (Ferrando et al., 2017) emphasize embedding the considerations of financial flexibility directly into experience and long-term investment measures. The study framework is structured and positionally suitable to develop into multi-year perspectives. Finally, in resonance to the recognition of a dynamic adaptation, (Carayannis et al., 2025) portray that embedding foresight and AI-based analytical measures make SMEs more resilient over time.

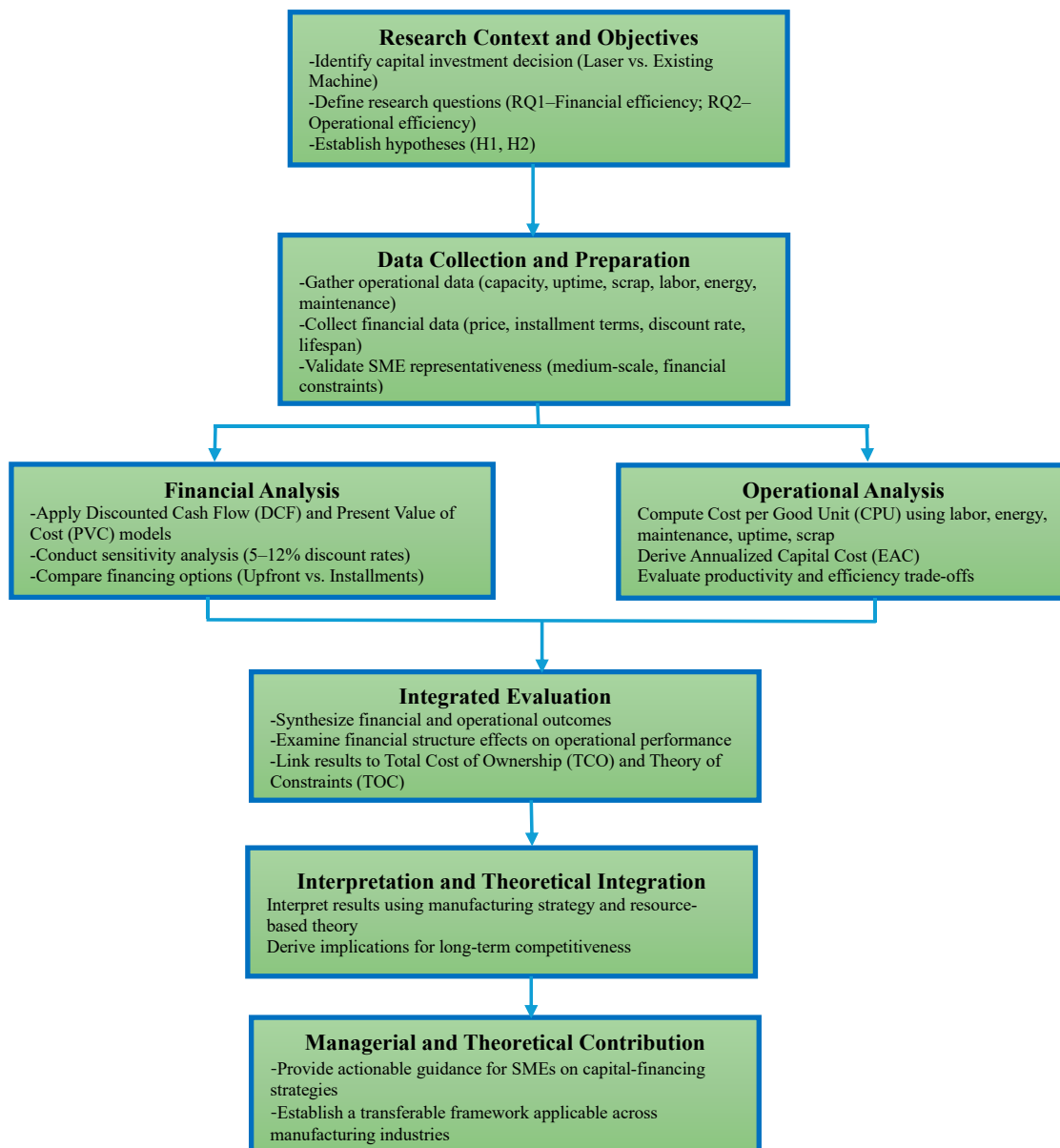
In summary, this research is methodologically robust and uniquely combines financial and operational analyses into a single

decision-making framework. The narrow case allows more specificity and insight into what is practical, but also more importantly it offers a route for broader validation. The two-year time horizon directs the focus on the immediate financial realities of SMEs. The authors also acknowledge uncertainty in developing long-term learning and strategic benefits, which allows for ongoing discovery. Consequently, it not only provided actionable advice, but has established a roadmap for continued innovation in manufacturing investment research.

3. Method

This study utilizes a case-based validation method to showcase a transferable financial and operational decision framework related to a manufacturing investment. The specific SME laser-cutting case is employed as an illustrative case study, its capital intensity and data transparency representative of many manufacturing contexts today. It is important to note that the objective here is a validation of the framework logic. It is not to address a real-world problem involving one firm to ensure sustained relevance to general manufacturing theory (Surma, 2015). As part of the analysis, the results were scrutinized in terms of two key issues (i) the financial consequences of the possible payment options and (ii) operational efficiency (productivity delivered by the new machine compared with the existing cutting machine). The overall research design incorporated a mixed method design, quantitative financial modelling and a more general analysis of production system performance. Figure 1 presents the research methodology.

Figure 1. Research development



3.1 Research Design

We undertook comparative evaluation research in two stages:

- Financial Evaluation: The total cost of ownership was evaluated by comparing the upfront payment option and installment financing option using the discounted cash flow approach, present value of cost (PVC) and some sensitivity analysis (Mellichamp, 2013).

- Operational efficiency Evaluation: For this analysis, we compared machine productivity including effective throughput and utilization rates, of the new versus the older/conventional cutting machine (Patrício et al., 2025).

Consequently, the empirical context should be understood as an example validation framework and not the focus of analysis. The practical purpose for developing the empirical case is to illustrate the analytical integration of standard measures of financial and operational activities to support decisions regarding investment in manufacturing, regardless of industry, country, or size. The empirical context allows for the clear identification and definition of the parameters of the framework but does not limit the overall applicability of the core principles of manufacturing reasoning.

3.2 Data Collection

Main data was collected via company visits as well as operational logs and interviews with management.

The following categories were examined:

Equipment data: Cycle time, speed, downtime, and utilization of both machines.

Machine (T= Old cutter machine) continuous operation with lower cycle speed

Machine (L = Laser cutting machine) long idle periods and high cycle speeds

Operation hours: 250 shifts/year \times 8 hours/day = 2000 scheduled hours yearly

Performance: capacity per shift, uptime ratio, and scrap rate

Cost data: maintenance, labor, energy, and machine purchase price

Financial data: The prices to purchase, installment terms (2 payments of \$500,000), discounted upfront cost (\$700,000), and current available discount rates 8% base case, lifespan = 5 years.

3.3 Context justification.

The SME chosen is a typical medium-scale discrete-manufacturing operation exhibiting constrained financial flexibility, and at best a moderate level of automation. The SME was selected with an emphasis upon access to in-depth operational and financial records that provide for an empirical context that may be representative of the purpose of validating the conceptual framework. The analysis will focus on methodological demonstration rather than sectoral description.

3.4 Financial Process Evaluations

The financial evaluations used discounted cash flow methods. The present value of cost (PVC) was calculated for each payment option as follows:

$$PVC = \sum (C_t / (1 + r)^t)$$

Where C_t is the cash outflow at time t , and r is the discount rate.

A univariate sensitivity analysis is performed on interest rates as we look to understand how the attractiveness of financing changes under varying cost of capital assumptions. The discount rate is varied from 5% to 12%, which provides a realistic picture of financing conditions.

3.5 Operational Process Evaluations

This study examined investment in a laser cutter through financial and operational assessments. Financially, present value of costs (PVC) analysis was used to compare upfront payment to installment payment for discount rates of (5%-12%). Operationally, cost per good unit was determined by combining production capacity, uptime, scrap, labor, energy, maintenance and annualized capital cost. Integrated accounting for both puts a realistic foundation for determining if the laser provides a commercial advantage relative to the current piece of equipment.

3.6 Reflecting on the Results

The evaluation included a multidimensional evaluative framework that incorporates both financial outcomes (PVC, total cost differences) as well as operational outcomes. As noted previously, it was important to also contextualize these findings by

incorporating strategy regarding liquidity issues, technological obsolescence, and competitive issues.

4. Results

The evaluation results will be provided in two parts: (i) financial performance analysis; and (ii) operational efficiency assessment. A sensitivity analysis is also provided to support the robustness of the financial results.

4.1 Analysis and Interpretation of Financial Performance

Two plans for financing were put forward. The following are the plans:

Option A (upfront): One-time payment of \$700,000 at the start of the project

Option B (Equated Installment Scheme): Payment of \$500,000 at the end of years 1 and 2

Basis/Boundary:

- Discount: $r=8\%$ per year
- Timeline: Two years

PVC Calculations:

Option A: $PVC_A = \$-700,000$

Option B: $PVC_B = -500,000 / (1.08)^1 - 500,000 / (1.08)^2$
 $= -462,963 - 428,669 = \$-891,632$

Comparison:

The difference in value: \$300,000 (\$1,000,000 - \$700,000)

The present value is now quantified to \$191,632

Option B cost is approximately 27.4% more than Option A. ($\$191,632 / \$700,000 = 0.2738 = 27.4\%$)

Sensitivity Analysis:

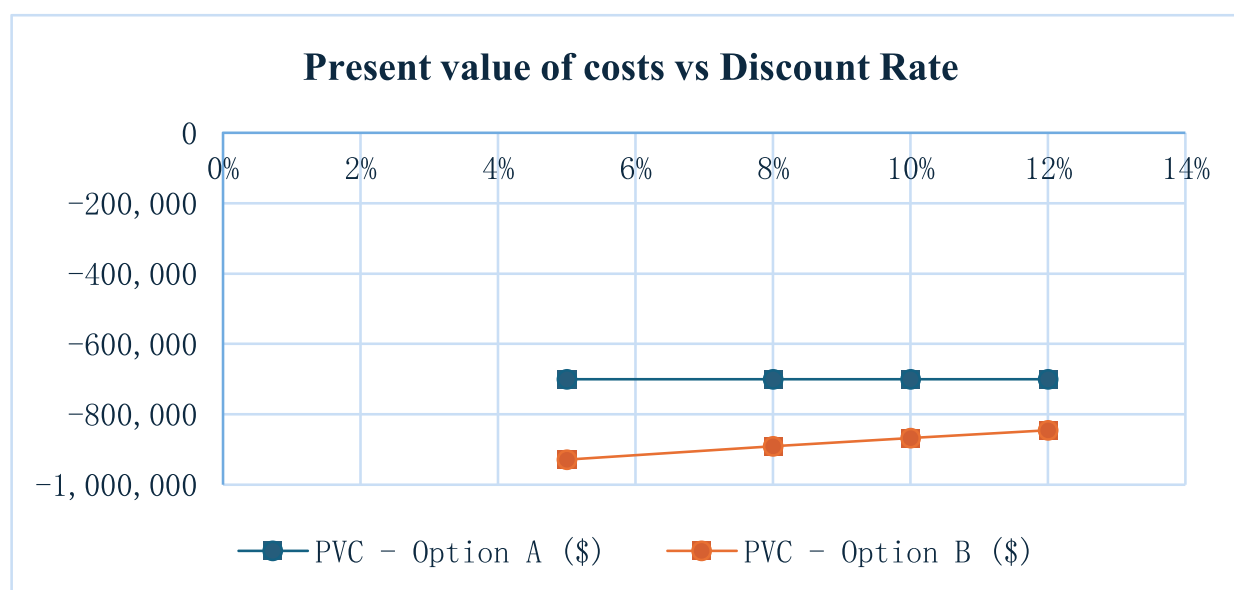
The present value of cost in Table 1 was recalculated using various discount rates (5 – 12%).

Table 1. Discount rates vs PVC options

Discount Rate	PVC - Option A (\$)	PVC - Option B (\$)	Difference (\$)	Preferred Option
5%	-700,000	-929,705	229,705	Upfront
8%	-700,000	-891,632	191,632	Upfront
10%	-700,000	-867,768	167,768	Upfront
12%	-700,000	-845,026	145,026	Upfront

The upfront option remains better in all cases. Figure 2 shows present value of costs vs discount rate.

Figure 2. Discount rate vs PVC for A and B



4.2 Operating efficiencies

Shifts/year = (50 weeks \times 5 days/week) = 250

Hours/shift = 8, scheduled annual hours = $250 \times 8 = 2,000$ h

Current machine (T)

Nameplate capacity per shift = 200 units

Uptime (availability) = 0.60

Scrap rate = 5%

Labor cost = \$25 / hour

Energy cost = \$10 / hour

Maintenance/year = \$50,000

Capital cost (PV) = \$400,000; (One payment)

Laser (L)

Nameplate capacity per shift = 100 units

Uptime = 0.95, Scrap rate = 2%

Labor cost = \$20 / hour

Energy cost = \$8 / hour

Maintenance/year = \$30,000

Capital PV (installment PV used as example) = \$891,632; life = 5 yrs; discount = 8%

Formulas used

Let:

L = Advanced laser cutter machine

T = Traditional cutter machine

H = scheduled annual hours = shifts/year \times hours/shift (here H = 2,000)

C = capacity per shift (units when running)

u = uptime (fraction)

s = scrap fraction

W = labor cost per hour

E = energy cost per hour

M = maintenance per year

K = annualized capital cost (Equivalent Annual Cost)

Produced units per year = $C \times \text{shifts/year} \times u$

Good units per year = Produced $\times (1 - s)$

Running hours per year = $H \times u$

Total annual cost = $(W + E) \times \text{running hours} + M + K$

Cost per good unit (CPU) = Total annual cost / Good units

Annualize capital (K) from PV:

$K = PV \times \text{annuity factor}$ where annuity factor = $r(1+r)^n / ((1+r)^n - 1)$

Base-case numeric result (with the assumptions above)

Annualized capital (5 years, $r=8\%$)

Annuity factor = 0.25045

$K_T = 400,000 \times 0.2504 = \$100,160/\text{year}$

$K_L = 891,632 \times 0.2504 = \$223,265/\text{year}$

Running-hours

T running hours = $2,000 \times 0.60 = 1,200$ h/yr

L running hours = $2,000 \times 0.95 = 1,900$ h/yr

Costs (annual), see Figure 3 for annual cost breakdown

$$T \text{ labor} = 25 \times 1,200 = \$30,000$$

$$T \text{ energy} = 10 \times 1,200 = \$12,000$$

$$T \text{ maintenance} = \$50,000$$

$$T \text{ capital annual} = \$100,160$$

$$T \text{ total annual cost} = \$192,160$$

$$L \text{ labor} = 20 \times 1,900 = \$38,000$$

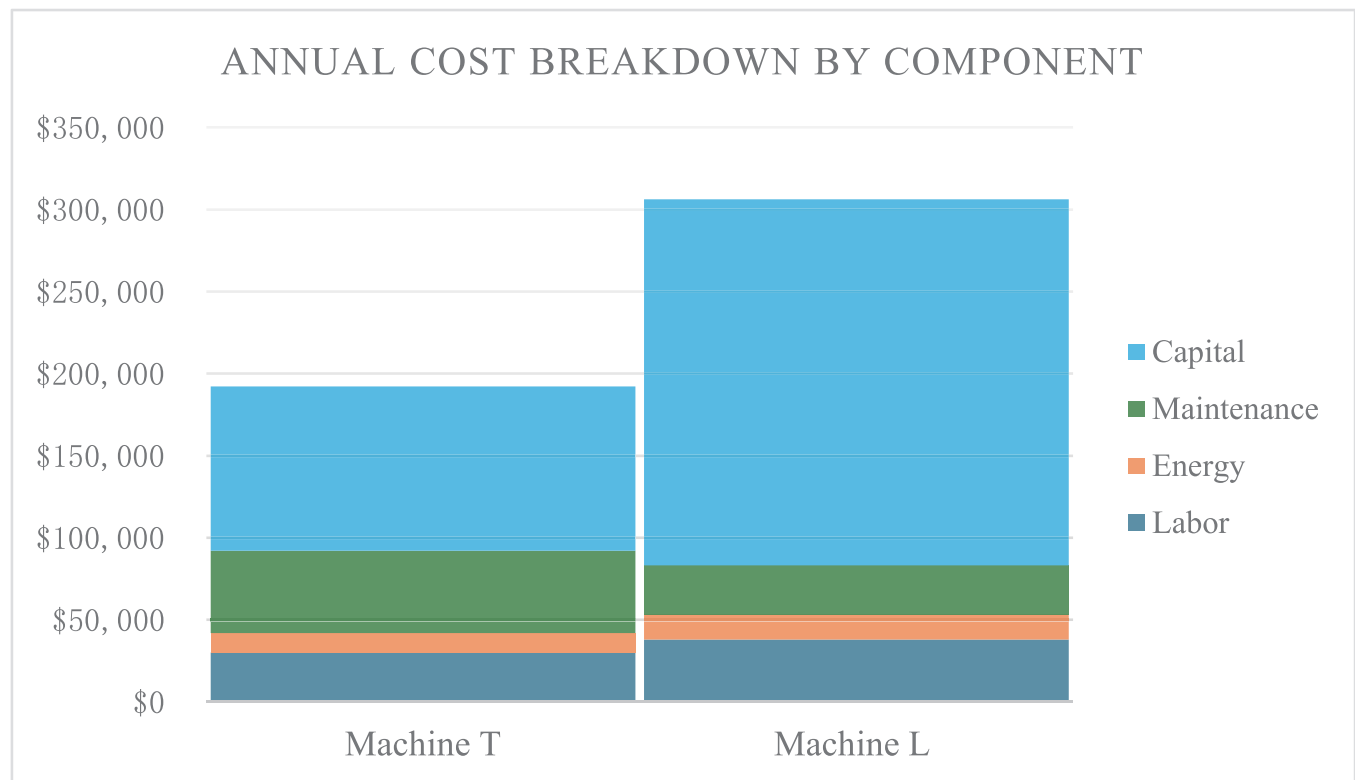
$$L \text{ energy} = 8 \times 1,900 = \$15,200$$

$$L \text{ maintenance} = \$30,000$$

$$L \text{ capital annual} = \$223,150$$

$$L \text{ total annual cost} \approx \$306,350$$

Figure 3. Annual cost breakdown



Output (annual)

$$T \text{ produced} = 200 \times 250 \times 0.60 = 30,000 \text{ units}$$

$$L \text{ produced} = 100 \times 250 \times 0.95 = 23,750 \text{ units}$$

Cost per good unit (CPU)

$$T \text{ good units} = 30,000 \times (1 - 0.05) = 28,500$$

$$\text{CPU}_T = \$192,160 / 28,500 = \$6.74 / \text{good unit}$$

$$L \text{ good units} = 23,750 \times (1 - 0.02) = 23,275$$

$$\text{CPU}_L = \$306,350 / 23,275 = \$13.16 / \text{good unit}$$

Despite higher uptime and lower scrap for the laser, the laser's higher capital (and lower nameplate capacity) produces a higher cost per good unit. With these inputs the laser is not economically justified on a per-unit cost basis. Figure 4 presents the cost per good unit for both machines, and Figure 5 presents an output comparison between old machines (T) and new laser machines (L).

Figure 4. Cost per good unit

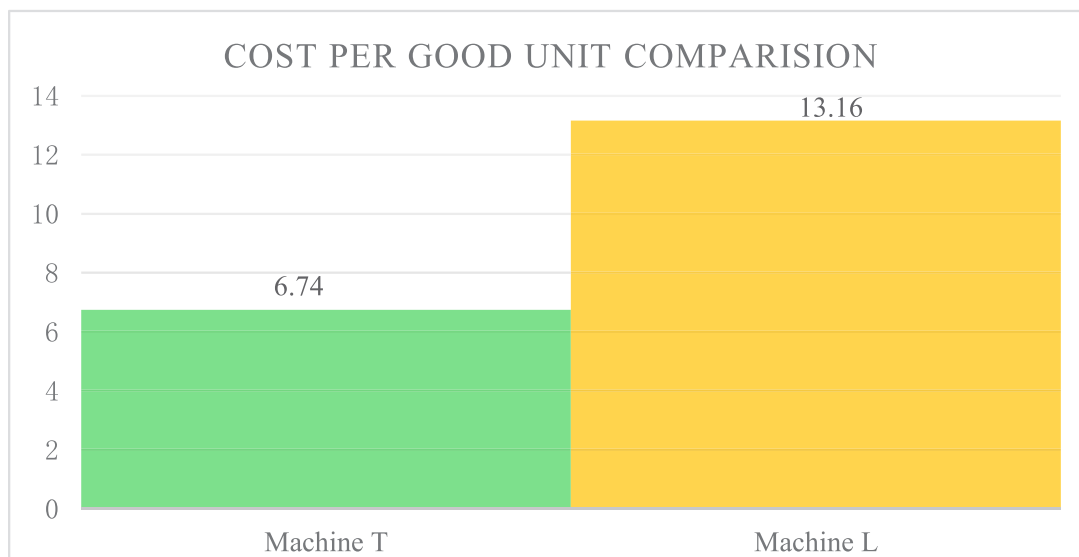
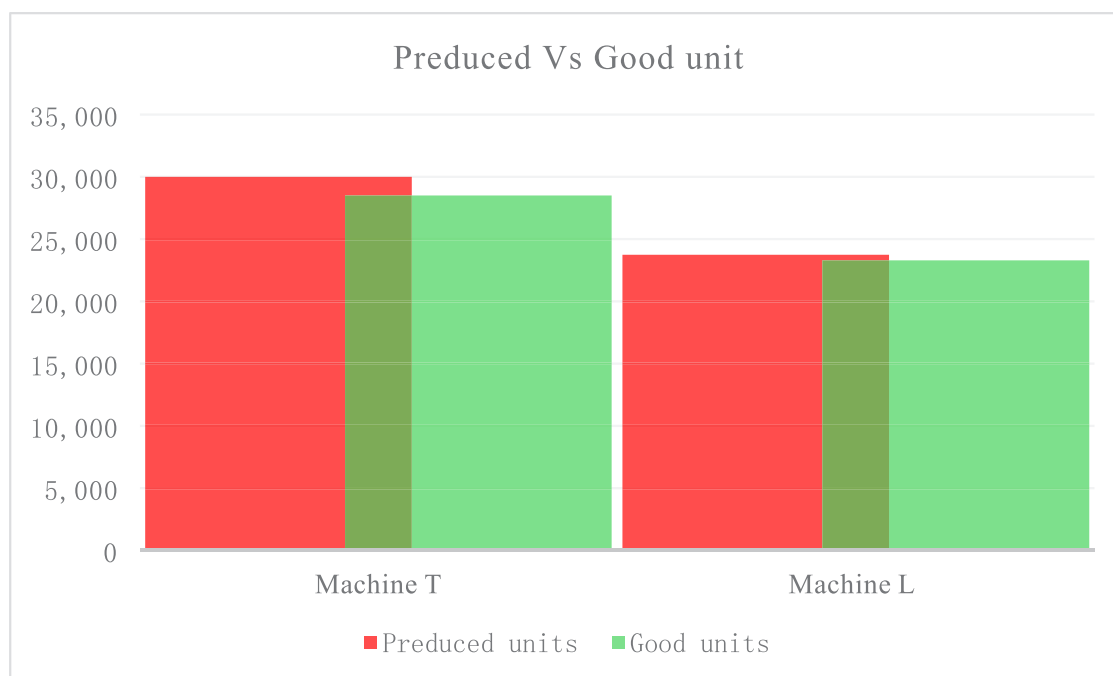


Figure 5. Output comparison



5. Discussion

In addition to the specific example used to illustrate this relationship, the results of the study demonstrate a general relationship between the capital structure of a manufacturer's financing and the operational cost effectiveness (OCE) of its product. Financing premium payments (through interest rates) are converted into annualized capital payment amounts, which then influence the cost structure of producing a good (i.e., cost per good unit) and may offset some of the perceived benefits of operational efficiencies (e.g. increased uptime or reduced scrap rates). This mechanism behaves consistently across all capital-intensive manufacturers and exemplifies an underlying characteristic of all capital-intensive manufacturers. The study's results support and expand the concept of a general theory for evaluating investments in the manufacturing sector.

This investigation examined whether the investment in a new laser cutting machine was economically viable, based upon two questions: (i) which payment plans are more cost-effective to the operation considering whatever assumptions we take regarding discount rates? and (ii) whether the operational performance of the laser provides value greater than its financial obligation (Shaw et al., 2004).

5.1 Financial performance

The PVC analysis clearly demonstrated that upfront payment was more financially advantageous than an installment plan (Mellichamp, 2013). The total present-value cost of the installment plan was higher than the upfront payment option across all discount rate scenarios ranging between 5% and 12%, with penalties of \$145k to over \$229k. Overall, installment financing, despite facilitating different cash flow timing, was a costly financing option due to the present-value cost of financing the premium. Accordingly, RQ1 is satisfied: upfront payment was a superior financing option under all circumstances tested. Furthermore, at 8% discount rate, the installment PV = \$891,632, which is \$191,632 (27.4%) more than the lump-sum option of \$700,000. A sensitivity analysis across 5% - 12% discount rates indicated that the PV for financing always exceeded the PV for lump-sum purchase (H1 satisfied).

5.2 Operational performance

Operational cost-per-unit analysis concluded the incumbent machine cost approximately \$6.7 per good unit while the laser was about \$13.2. The difference in cost is linked in the main to the capital charge associated with the laser machine, even though it had lower nominal capacity, which was substantially larger than operational uptime (95% to 60%) and lower scrap rates (2% to 5%) for laser cut machine. So, we can consider RQ2 resolved: while functional productivity obtained with the laser cutting machine was superior to the incumbent, the financial premium associated with financing a laser yielded no value for productivity gains (Baumers et al., 2016). The incumbent technology's CPU = \$6.74 vs. laser CPU = \$13.16 indicating that the incumbent was actually almost twice as cost-effective given the actual conditions observed. Even assuming a higher uptime (95% vs. 60%) and throughput with the laser, the financing and any scrap or maintenance costs outweigh these increases and were more than adequate to account for the higher capacity (H2 confirmed).

5.3 Interpretation using theory

Since the experiments had outcomes consistent with several different theoretical frameworks, we will interpret the results through each theory separately. The results follow a Total Cost of Ownership (TCO) logic (Rahman, 2025 ; Ellram, 1993): in analyzing alternatives the operating, capital and quality cost or costs, must be considered. Our analysis was also consistent with and incorporates concepts from the Theory of Constraints (TOC) literature: enhancing the reliability of one machine will not necessarily increase throughput if another machine is the bottleneck, in this experiment this is typified by the nominal lower throughput of the laser cutting machine. Finally, while we did not calibrate with respect to the ambience repayments and learning dynamics of new technology explicitly, the findings do seem consistent with a real-options perspective on learning; once with the uncertainty resolved (Johnson & Scott, 2019; Foss et al., 2025), and their capital risk within acceptable limits, many organizations will value new technology all the more, e.g., replacing a cutter may only bring payback or value in operation was evident when certain procurement practices such as service contracts or staged adoption may reduce potential risks.

In addition to verifying the consistency of theoretical construction, this research strengthens the relationship between firms' manufacturing strategy and their resource-based view of competitiveness. The findings indicate that firms' resources (i.e., the combination of financial flexibility and operational capability) jointly inform their competitive position as described in the resource-based view of competitiveness. The model demonstrates how the choice of financial arrangements affects production costs and systems' level of efficiencies; in addition, it shows how the decisions related to a firm's capital structure will subsequently impact a firm's long-term competitiveness and its ability to respond to environmental changes within a technology-intensive manufacturing environment.

5.4 Contribution to manufacturing theory

The results contribute to manufacturing theory by connecting financial structure and operational performance via a single evaluative construct. The results reveal that financing structure (installments versus lump sum) can alter annualized capital loadings and therefore change apparent productivity benefits, a mechanism that has been ignored in research on manufacturing systems. Thus, the framework extends total-cost-of-ownership and technology-adoption models to incorporate the finance-operations interface and provides propositions relevant in both manufacturing contexts, where capital is relatively fixed and units can only be utilized at limited efficiencies.

These observations, although based on a single case, indicate structural relationships between financing form, cost of capital, and operational performance that are common to all manufacturing systems. Therefore, this conversation now shifts from a case-specific consideration to more general management and theoretical implications.

5.5 Implications for practice

The findings have three implications for managers. First, they should care about how they finance the machine: a lease-purchase or instalment line may hide a significant resulting cost burden on your margins, for the most part paying an outright price for it up front is the least transactional cost and of least uncertainty. Second, they should care about the terms of procurement: the anonymous breakeven analysis (this analysis is part of future work) provides an objective capital reduction development goal (important for the contracting authority) for estimating a competitive position (i.e. where their commercial position will be with respect to TCO) to be better than the shaded table provided procurement negotiators or pay more if a unit. Third, they should have a system view: Unless the cutter is the bottleneck of the value stream, replacing it does not necessarily model or id SC spending, for either throughput or unit cost of goods sold irrespective of how much better its uptime or deal for “punctuate” costs.

The study expands on existing theories regarding manufacturing investment by formally integrating the financial structure and operational performance together using one unified evaluative measure. In addition, instead of treating capital budgeting as two independent domains from the realm of manufacturing efficiency, the framework helps clarify how financial decisions can flow into the manufacturing system, ultimately affecting the productivity results within the manufacturing system. This type of integration helps to push forward the development of manufacturing theory by showing a linkage between the financial aspects of the business and the operational aspects of the manufacturing business, which is a relatively unexplored area of research.

5.6 Implications for research

This study adds to the literature by combining discounted financial evaluation with operational cost-per-unit evaluation and analyzing the interaction between the two dimensions of analysis (Hobdari et al., 2009; Ramasesh & Jayakumar, 1997). In the past, this interaction aspect has often been ignored or treated as separate, while our findings illustrate the need to recognize the links between financial and operational issues, particularly for SMEs, where constraints on available capital, along with operational variability, are strongly associated with the investment decision. Future research has the potential to extend this line of analysis to multiple stages instead of simply a ‘one-off’; to integrate learning effects in a more explicit manner than has happened in prior work; and to investigate alternative forms of procurement such as leasing or performance-based contracts (Glas & Kleemann, 2017).

The research is based on an industrial case with a specific machine and a specific financing arrangement; however, it is much more than a single application. The integrated framework indicates that manufacturers can expand their decision criteria from just financial to include operational efficiency for capital investment. By combining discounted cash flows with a “cost-per-good-unit” measure of productivity, the research provides a transferable decision-support tool that can help firms in other industries align equipment purchases beyond financial considerations and towards operational efficiency and long-term production objectives.

Overall, the convergence of the financial and operational evidence suggests that, at this time, the incumbent machine is the economically rational choice. The laser will only make sense if either the overall capital costs can get so low that it becomes attractive, or if there is some other benefit, beyond unit cost, that is important to the larger strategic vision, such as new capability in products, or differentiation through quality.

6. Conclusion

While the framework and findings of this study have been empirically tested using a case study of a single producing business, the same scenario could apply across all different types of production/buying decisions. And therefore, the recommendations and insights from this framework may be very valuable to a substantial number of manufacturing companies that have significant capital investments and financing constraints impacting their capital investment decisions.

The study used financial and operational analyses to consider whether to adopt a laser cutting machine in lieu of the

incumbent cutter. While financial analysis using discounted cash flow indicated that upfront financing offers a discount rate benefit over installment payments, upfront financing provided an overall cost savings of \$191,632 (27.4%) at an 8% discount rate. While operational analysis showed the laser improves downtime (95% vs 60%), with half the scrap rate (2% compared to 5%) of the incumbent cutting machine, greater downtime and the significant differences in annualized capital costs resulted in a total cost per good unit nearly doubling within the incumbent (\$13.16 compared to \$6.74).

In summary, the findings show that the existing machinery is the better decision if current costs do not change. We have presented a way to examine the combination of Total Cost of Ownership (TCO) and operational productivity and demonstrated that equipment investment decisions need to be made in the context of system constraints. The study also highlighted the potential for managers to wrongly allocate their attention to the performance of equipment without consideration of financial constraints. It is acceptable for managers to use financial representations, however, these should be used and understood within a complete operational cost analysis.

Future works could also consider a break-even assessment to identify when the laser cutter becomes competitively priced; potentially applying other financing structures and discount rates. Within the current scope of the research project, the framework can be modified to suit other manufacturing environments and provide decision-makers with a usable framework for considering how to balance financial cost to operational efficiency.

Theoretical contribution and generalizability. Although validated in one SME, the integrated framework could be transferred to almost any manufacturing context where equipment-investment trade-offs are present. By decontextualizing from the particular case, it provides a unified theory of manufacturing investment evaluation that bridges financial discipline with operational capability. The combined model is an example of high flexibility from the manufacturing environment in which it was developed, with properties and performance measures transferable to equipment decision making across manufacturing sectors, i.e. automotive, aviation, electronics or large/medium equipment manufacturing. Therefore, the model is a general tool that can be leveraged to support decision making regarding investments associated with Industry 4.0.

Funding

No

Conflict of Interests

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

Reference

- [1] Rahman, M. M. (2025). Data analytics for strategic business development: A systematic review analyzing its role in informing decisions, optimizing processes, and driving growth. *Journal of Sustainable Development Policy*, 1(1), 285–314. <https://doi.org/10.63125/hel1tfg25>
- [2] Amjad, S., Neelakrishnan, S., & Rudramoorthy, R. (2010). Review of design considerations and technological challenges for successful development and deployment of plug-in hybrid electric vehicles. *Renewable and Sustainable Energy Reviews*, 14(3), 1104–1110. <https://doi.org/10.1016/j.rser.2009.11.001>
- [3] Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. Boston Consulting Group, 9(1), 54–89.
- [4] Lanteri, A., & Rampini, A. A. (2025). Financing the adoption of clean technology (Working Paper No. w33545). National Bureau of Economic Research.
- [5] Hobdari, B., Jones, D. C., & Mygind, N. (2009). Capital investment and determinants of financial constraints in Estonia. *Economic Systems*, 33(4), 344–359. <https://doi.org/10.1016/j.ecosys.2009.05.004>
- [6] Zeidan, R., & Shapir, O. M. (2017). Cash conversion cycle and value-enhancing operations: Theory and evidence for a free lunch. *Journal of Corporate Finance*, 45, 203–219. <https://doi.org/10.1016/j.jcorpfin.2017.04.014>
- [7] Baumers, M., Dickens, P., Tuck, C., & Hague, R. (2016). The cost of additive manufacturing: Machine productivity, economies of scale and technology-push. *Technological Forecasting and Social Change*, 102, 193–201. <https://doi.org/10.1016/j.techfore.2015.02.015>

- [8] Driouchi, T., & Bennett, D. J. (2012). Real options in management and organizational strategy: A review of decision-making and performance implications. *International Journal of Management Reviews*. <https://doi.org/10.1111/j.1468-2370.2011.00304.x>
- [9] Johnson, B. A., & Scott, H. S. (2019). Controlling the long-term problem of short-term funding. *Journal of Financial Regulation*, 5(2), 101–149. <https://doi.org/10.1093/jfr/fjz004>
- [10] Burlea-Schiopoiu, A., & Mihai, L. S. (2019). An integrated framework on the sustainability of SMEs. *Sustainability*, 11(21), 6026. <https://doi.org/10.3390/su11216026>
- [11] Carayannis, E. G., Dumitrescu, R., Falkowski, T., Papamichail, G., & Zota, N. R. (2025). Enhancing SME resilience through artificial intelligence and strategic foresight: A framework for sustainable competitiveness. *Technology in Society*, 81, 102835. <https://doi.org/10.1016/j.techfore.2025.122586>
- [12] Ellram, L. M. (1993). A framework for total cost of ownership. *International Journal of Logistics Management*, 4(2), 49–60. <https://doi.org/10.1108/09574099310804984>
- [13] Ferrando, A., Marchica, M. T., & Mura, R. (2017). Financial flexibility and investment ability across the Euro area and the UK. *European Financial Management*, 1, 87–126. <https://doi.org/10.1111/eufm.12091>
- [14] Foss, N. J., Klein, P. G., & Murtinu, S. (2025). Entrepreneurial judgment, uncertainty, and resource mobilization. *Review of Austrian Economics*, 1–9. <https://doi.org/10.1007/s11138-025-00690-6>
- [15] Gherghina, Ș. C., Botezatu, M. A., Hosszu, A., & Simionescu, L. N. (2020). Small and medium-sized enterprises (SMEs): The engine of economic growth through investments and innovation. *Sustainability*, 12(1), 347. <https://doi.org/10.3390/su12010347>
- [16] Glas, A. H., & Kleemann, F. C. (2017). Performance-based contracting: Contextual factors and the degree of buyer–supplier integration. *Journal of Business & Industrial Marketing*, 32(5), 677–692. <https://doi.org/10.1108/JBIM-04-2016-0065>
- [17] Markus, G., & Rideg, A. (2021). Understanding the connection between SMEs’ competitiveness and cash flow generation: An empirical analysis from Hungary. *Competitiveness Review*, 31(3), 397–419. <https://doi.org/10.1108/CR-01-2020-0019>
- [18] Mellichamp, D. A. (2013). New discounted cash flow method: Estimating plant profitability at the conceptual design level while compensating for business risk/uncertainty. *Computers & Chemical Engineering*, 48, 251–263. <https://doi.org/10.1016/j.compchemeng.2012.08.012>
- [19] Patrício, L., Varela, L., & Silveira, Z. (2025). Proposal for a sustainable model for integrating robotic process automation and machine learning in failure prediction and operational efficiency in predictive maintenance. *Applied Sciences*, 15(2), 854. <https://doi.org/10.3390/app15020854>
- [20] Ramasesh, R. V., & Jayakumar, M. D. (1997). Inclusion of flexibility benefits in discounted cash flow analyses for investment evaluation: A simulation/optimization model. *European Journal of Operational Research*, 102(1), 124–141. [https://doi.org/10.1016/S0377-2217\(97\)00223-3](https://doi.org/10.1016/S0377-2217(97)00223-3)
- [21] Shaw, M., Williams, W., House, R., & Haynam, C. (2004). Laser performance operations model. *Optical Engineering*, 43(12), 2885–2895. <https://doi.org/10.1117/1.1815004>
- [22] Surma, J. (2015). Case-based approach for supporting strategy decision making. *Expert Systems*, 4, 546–554. <https://doi.org/10.1111/exsy.12003>

Why Do Graduates Turn Away from Private Firms? Examining the Roles of Perceived Job Security, Perceived Management Compliance, Perceived Humanistic Care, and Organisational Attractiveness in the Chinese Context

Weixiang Gan¹, Mengfei Xiao*, Naiqian Zhang, Yongli Zhang, Tara Ahmed Mohammed

1. Zhenguigu Technology Co., Ltd., Guangyang District, Langfang, Hebei 065000, China

2. Graduate school of Business, SEGi University, Kota Damansara, Petaling Jaya 47810, Selangor, Malaysia

3. Chongqing University of Arts and Sciences, Yongchuan District, Chongqing 402160, China

4. Graduate School of Business, SEGi University, Kota Damansara, Petaling Jaya 47810, Selangor, Malaysia

5. Graduate School of Business, SEGi University, Kota Damansara, Petaling Jaya 47810, Selangor, Malaysia

**Corresponding author: Mengfei Xiao, 282584787@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: At a time when private enterprises have become the principal absorbers of university graduates in China, an increasing number of young people continue to place their career aspirations in the public sector. This contrast has emerged as a salient contradiction within China's employment structure. Focusing on the question of why graduates avoid private enterprises, this study develops a research model grounded in signalling theory, incorporating perceived humanistic care, perceived management compliance, perceived job security, organisational attractiveness, and intention to seek employment in private enterprises. The model is empirically tested using data collected from 446 Chinese final-year university students and recent graduates within two years of graduation. The findings indicate that all three forms of perception significantly enhance graduates' intentions to pursue employment in private enterprises, and that each exerts its influence through organisational attractiveness as a mediating mechanism. Among them, perceived job security has the strongest effect, followed by perceived management compliance, while perceived humanistic care also plays a meaningful role and should not be overlooked. The results suggest that graduates are not inherently resistant to private enterprises; rather, what they seek to avoid is a career starting point characterised by uncertainty, disorder, and a lack of respect. This study not only sheds light on the formation mechanism of employment intentions towards private enterprises in the Chinese context, but also offers a more targeted explanatory framework and practical guidance for enhancing the attractiveness of private firms to young talent.

Keywords: Intention to Seek Employment in Private Enterprises; Recent Graduates; Signalling Theory; Corporate Humanistic Care; Job Security

Published: Mar 29, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1193>

Preface

This paper is dedicated to all university students in China who continue to smile despite struggling within an intensely

competitive and exhausting job market.

You are hardworking and self-assured, though at times the accelerating pace of the era may leave you feeling overwhelmed and breathless. Even so, I hope you will continue to believe in yourselves and trust that life will, in time, reward wisdom with the happiness it deserves. There is no need to take off Kong Yiji's long gown; it is your own battle robe.

This paper is written for you, with the hope that one day it may, in some way, serve you.

1. Research Background

Private enterprises in China have become one of the largest sources of employment for young people. According to publicly available data released by the National Development and Reform Commission, by the end of May 2025, the number of registered private economic entities in China had reached 185 million, accounting for 96.76% of all market entities nationwide. The private sector contributes more than 50% of national tax revenue, over 60% of gross domestic product, more than 70% of technological innovation outcomes, and over 80% of urban employment, while also encompassing more than 92% of the country's national high-tech enterprises (National Development and Reform Commission [NDRC], 2025). In addition, the All-China Federation of Industry and Commerce noted in its 2025 report on the “Hundred Cities, Thousand Schools, Ten Thousand Enterprises” employment promotion initiative that, over the past five years, graduates entering private enterprises have accounted for more than half of those whose employment destinations had been confirmed across various localities (All-China Federation of Industry and Commerce [ACFIC], 2025). These figures indicate that private enterprises are not merely a supplementary component of China's employment system; rather, they constitute one of the most important employment destinations for university graduates. The difficulty, however, is that their capacity to absorb labour in quantitative terms has not naturally translated into a preferred career option in the minds of graduates. For many young job seekers, private enterprises are often associated with fast-paced work routines, relatively small organisational scale, greater flexibility in rule implementation, and uncertain prospects for job retention. As a result, although private firms offer a large number of vacancies, they do not necessarily possess a level of talent attractiveness commensurate with the volume of positions available.

The continued expansion in the supply of university graduates has made this divergence in employment preferences increasingly visible. The Ministry of Education disclosed that the number of university graduates in the Class of 2025 was expected to reach 12.22 million, an increase of 430,000 over the previous year, marking the continuation of a multi-year trend in which graduate numbers have remained above the ten-million threshold (Ministry of Education of the People's Republic of China [MOE], 2024a, 2024b). As the graduate population continues to grow, employment choices have not dispersed evenly across different types of organisations; instead, a pronounced preference for the public sector has emerged. An analytical article published in 2026 by the Beijing Municipal Human Resources and Social Security Bureau, drawing on Zhaopin's 2024 annual survey, reported that the proportion of fresh graduates aspiring to enter state-owned organisations had risen from 47.9% in 2018 to 73.1% in 2024, whereas the proportion choosing private enterprises had fallen from 20.9% to 12.5% over the same period (Beijing Municipal Human Resources and Social Security Bureau, 2026). In its deployment plan for graduate employment for the Class of 2025, the Ministry of Education called for greater efforts to expand market-oriented and socially driven employment channels, to deepen the implementation of the “Hundred Cities, Thousand Schools, Ten Thousand Enterprises” initiative, and to increase job opportunities in private enterprises and small and medium-sized firms (MOE, 2024b). This policy expression indirectly suggests a substantial gap between the distribution of graduate destinations and the distribution of employment opportunities in the labour market. Put differently, what stands before fresh graduates is not simply the question of whether jobs exist, but whether they are willing to place the starting point of their careers in private enterprises.

2. Research Background

This study seeks to address the relatively weak willingness of Chinese fresh graduates to pursue employment in private enterprises by identifying the specific factors that influence whether private firms are considered a viable starting point for their careers, and by further examining how these factors enter graduates' job choice evaluations through organisational

attractiveness. In its employment deployment plan for the Class of 2025, the Ministry of Education explicitly called for stronger support for small and medium-sized enterprises in absorbing employment, the implementation of targeted recruitment initiatives encouraging private enterprises to hire university graduates, and the encouragement of higher education institutions across different regions to proactively align with the talent needs of leading technology firms, gazelle firms, and specialised and sophisticated SMEs that produce novel and unique products (Ministry of Education of the People's Republic of China [MOE], 2024). Subsequently, the Ministry of Human Resources and Social Security, the Ministry of Education, and the Ministry of Finance jointly issued a notice calling for the expansion of enterprises as the primary channel for employment and for the continued organisation of initiatives such as Private Enterprise Recruitment Month, city-level joint recruitment programmes, and online campus recruitment activities, with the aim of promoting a fuller transition of university graduates into market-oriented positions (Ministry of Human Resources and Social Security of the People's Republic of China et al., 2025). Against this policy backdrop, the present study aims to identify, through empirical investigation, the conditions under which fresh graduates' acceptance of and willingness to choose private enterprises may be strengthened, thereby providing evidence to help alleviate the excessive concentration of graduate destinations in public-sector institutions and to promote a more balanced distribution of employment choices.

3. Hypotheses Development

For fresh graduates, whether a private enterprise is considered a viable employment option often depends first on how they imagine they would be treated after entering the organisation. Signalling theory in recruitment research suggests that job seekers do not have direct access to the internal workings of an organisation and therefore rely on external cues released by the employer to infer what their future work experience may be like. Among these cues, expectations regarding whether employees will be treated with dignity, respect, and fairness are especially influential in shaping applicants' willingness to move closer to an organisation (Jones et al., 2014; Jones et al., 2016). This line of reasoning closely corresponds to the psychological judgment captured by perceived humanistic care. If fresh graduates believe that a private enterprise places greater value on individual feelings in its organisational climate, employee treatment, and day-to-day management practices, and is more inclined to treat newcomers with respect rather than in a draining or instrumental manner, they are more likely to regard that enterprise as an appropriate place to begin their careers. Research on Generation Z job seekers likewise shows that young applicants are concerned not only with pay and benefits, but also with workplace atmosphere, the sincerity of interpersonal relations, and whether the organisation conveys an ethical and humane way of treating people. Such non-material attributes enter directly into the formation of employment intentions (Nguyen Ngoc et al., 2022). On this basis, the following hypothesis is proposed:

H1: Perceived humanistic care positively influences Chinese fresh graduates' intention to seek employment in private enterprises.

Beyond whether an organisation appears caring, fresh graduates also assess whether a private enterprise operates according to clear and acceptable rules. Applicant reaction research has long indicated that job seekers form an overall evaluation of organisational fairness on the basis of procedures used in recruitment, selection, and management. Procedural justice, information transparency, and consistency in rule implementation all enter into their assessment of organisational attractiveness and job pursuit intentions (Gilliland, 1993). Subsequent studies have further shown a stable relationship between perceptions of organisational justice and applicants' attractiveness evaluations, with procedural fairness in particular being clearly associated with whether job seekers are willing to approach a given organisation (Bustaman et al., 2020). In the context of job seeking in Chinese private enterprises, perceived management compliance should not be understood narrowly as legal compliance alone. Rather, it refers to graduates' overall judgment regarding whether the enterprise manages people in a standardised manner, whether institutional boundaries are clear, and whether managerial behaviour is applied consistently. For fresh graduates who are only beginning to enter the labour market, a private enterprise that is seen as having clear systems, non-arbitrary management, and recruitment and employment practices that are less dependent on personal preference is likely to be regarded as more credible and, therefore, more likely to be included in their set of preferred

employment options. On this basis, the following hypothesis is proposed:

H2: Perceived management compliance positively influences Chinese fresh graduates' intention to seek employment in private enterprises.

The choice of a career starting point is also closely related to applicants' assessment of job retention risk and their sense of occupational stability. Research on employer attractiveness and job choice preferences has repeatedly noted that organisational stability, job security, and confidence in continued employment are among the organisational attributes frequently considered by job seekers, particularly those entering the labour market for the first time (Jain & Bhatt, 2015; Mauger & Bryant-Lees, 2022). Studies focusing on university students and young job seekers further indicate that job security, together with benefits, career development, and organisational reputation, is commonly treated as one of the key criteria for judging whether an employer is worth joining (Kumari & Saini, 2018). In the context of Chinese fresh graduates considering employment in private enterprises, perceived job security extends beyond a narrow concern about layoffs. It also encompasses whether the position is likely to be interrupted, whether the beginning of one's career is likely to be marked by repeated job-hopping, and whether entry into the enterprise can provide a relatively stable occupational foothold. If graduates believe that a private enterprise can offer a reassuring expectation of continued employment, their resistance towards such firms is likely to decline, while their willingness to apply for and accept employment is more likely to increase. On this basis, the following hypothesis is proposed:

H3: Perceived job security positively influences Chinese fresh graduates' intention to seek employment in private enterprises.

Organisational attractiveness provides a psychological transmission mechanism through which the above three forms of perception are translated into employment intentions. Recruitment research generally suggests that job seekers do not immediately decide whether to apply to or join an organisation upon receiving cues about its attributes. Rather, they first develop an overall judgment as to whether the organisation is worth approaching, and this general evaluation is commonly conceptualised as organisational attractiveness. Relevant studies indicate that applicants' interpretations of organisational signals concerning respect, fairness, values, and working conditions first shape their evaluation of organisational attractiveness, which then further influences job pursuit intentions and application behaviour (Ahamad et al., 2023; Khan & Muktar, 2020). In other words, fresh graduates may become more willing to choose a particular private enterprise not simply because they perceive one isolated positive characteristic, but because these favourable attributes accumulate into an overall sense of liking and willingness to approach the organisation, which is then converted into a clearer employment preference. Following this logic, perceived humanistic care, perceived management compliance, and perceived job security are expected not only to enhance graduates' perceptions of organisational attractiveness towards private enterprises, but also for organisational attractiveness itself to increase their intention to seek employment in such firms and to mediate the relationships between the three perceptions and employment intention. On this basis, the following hypotheses are proposed:

H4: Perceived humanistic care positively influences organisational attractiveness.

H5: Perceived management compliance positively influences organisational attractiveness.

H6: Perceived job security positively influences organisational attractiveness.

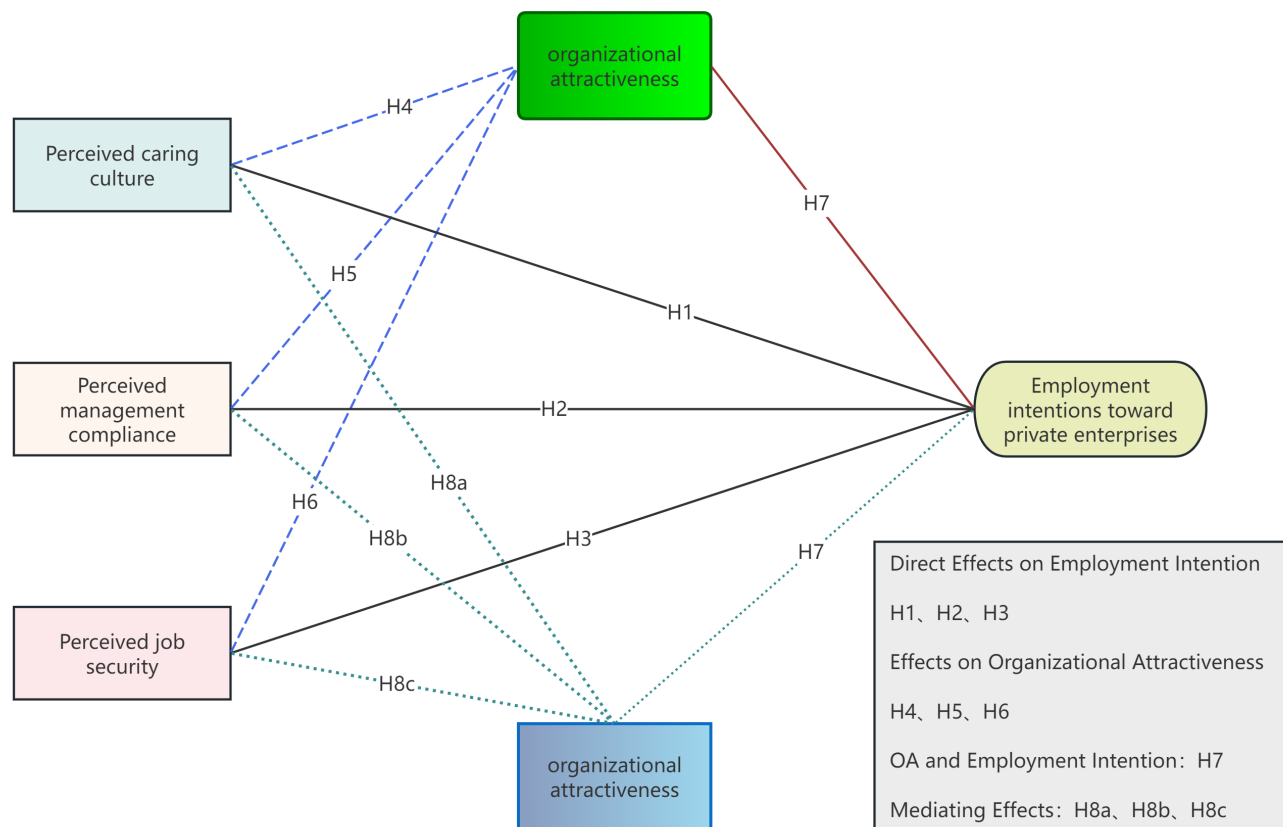
H7: Organisational attractiveness positively influences Chinese fresh graduates' intention to seek employment in private enterprises.

H8a: Organisational attractiveness mediates the relationship between perceived humanistic care and Chinese fresh graduates' intention to seek employment in private enterprises.

H8b: Organisational attractiveness mediates the relationship between perceived management compliance and Chinese fresh graduates' intention to seek employment in private enterprises.

H8c: Organisational attractiveness mediates the relationship between perceived job security and Chinese fresh graduates' intention to seek employment in private enterprises.

Graphic 1. Conceptual Framework



4. Research Design

This study adopts a quantitative approach and is conducted within a deductive research design. This choice is warranted because the core constructs examined in this study, namely perceived humanistic care, perceived management compliance, perceived job security, organisational attractiveness, and intention to seek employment in private enterprises, all pertain to psychological cognition and attitudinal judgment. Unlike demographic characteristics such as age, income, or educational attainment, these constructs cannot be directly observed. They can only be captured indirectly through multiple measurement items designed to reflect the underlying latent variables, after which statistical modelling can be employed to test the structural relationships among them. Research on job seekers' evaluations of organisations and their employment intentions has long relied on questionnaire-based methods to collect latent construct data and on structural equation modelling to examine multivariate relationships, precisely because such constructs are fundamentally grounded in individuals' subjective perceptions and psychological representations (Highhouse et al., 2003; Jones et al., 2016). The theoretical propositions and chain of hypotheses guiding this study have already been established in the preceding sections. Accordingly, the purpose of the present research is not to generate concepts inductively from field data, but to subject theoretically derived propositions to empirical testing on the basis of existing literature. Given that the proposed model incorporates multiple latent constructs, several direct paths, and mediating relationships, the subsequent data analysis will be carried out using structural equation modelling. Methodological studies have shown that PLS-SEM is comparatively less restrictive with respect to sample size and distributional assumptions, making it particularly suitable for analysing path models composed of multiple latent variables (Dash & Paul, 2021; Hair et al., 2021). For this reason, PLS-SEM is adopted in the present study.

The questionnaire design is grounded in construct dimensions rather than being developed intuitively or item by item without theoretical structure. Perceived humanistic care is conceptualised from three interrelated aspects: respectful treatment, caring support, and a sense of dignity in everyday interactions. Its theoretical foundation draws, first, on research concerning workplace dignity and respectful interaction, which emphasises equality, respect, being valued, and being recognised in the work experience, and second, on studies of perceived organisational support, which focus on whether the organisation

is concerned with members' well-being and needs (Thomas & Lucas, 2019; Kurtessis et al., 2017). Perceived management compliance is measured through several dimensions, including procedural consistency, information transparency, rule clarity, and perceived managerial fairness, primarily based on applicant reaction and organisational justice research addressing procedural fairness, adequacy of explanation, and the predictability of organisational rules (Gilliland, 1993; Truxillo et al., 2004). Perceived job security is designed around expected continuity of employment, confidence in job retention, clarity of the employment relationship, and perceived risk of unemployment, drawing on studies of employment security rights and employer attractiveness (Roehling et al., 2000; Vogel & Satzger, 2024). Organisational attractiveness is measured following the tradition established by Highhouse et al., with an emphasis on respondents' overall evaluations of whether the enterprise is worth approaching, whether it could be regarded as an ideal employer, and whether they would be willing to learn more about it or move closer to it as a career option (Highhouse et al., 2003). The dependent variable, intention to seek employment in private enterprises, is developed with reference to research on job pursuit intentions and application intentions, with items covering willingness to submit an application, willingness to continue in the recruitment process, willingness to accept an offer, and willingness to include the enterprise among one's preferred employment options (Khan & Muktar, 2020; Nguyen Ngoc et al., 2022). At the operational level, each construct is measured using four to five items on a five-point Likert scale, where 1 indicates "strongly disagree" and 5 indicates "strongly agree". The item development process begins with the extraction of relevant wording from established scales, followed by bilingual adaptation and semantic condensation to fit the specific context of Chinese fresh graduates seeking employment in private enterprises. The resulting items are then reviewed by experts in management and human resource research to ensure content relevance and clarity of expression.

Table 1. Basis for Determining the Construct Dimensions of Each Variable

Variable	Construct Dimension	References
Perceived Humanistic Care (PHC)	Respectful Treatment	Thomas and Lucas (2019) Kurtessis et al. (2017)
	Caring Support	
	Sense of Dignity in Daily Interactions	
Perceived Management Compliance (PMC)	Procedural Consistency	Gilliland (1993) Truxillo et al. (2004)
	Information Transparency	
	Rule Clarity	
	Perceived Managerial Fairness	
Perceived Employment Security (PES)	Expected Job Continuity	Roehling et al. (2000) Vogel and Satzger (2024)
	Confidence in Retention	
	Clarity of Employment Relationship	
	Perceived Risk of Job Loss	
Organisational Attractiveness (OA)	Overall Evaluation of Worth Approaching	1. Highhouse et al. (2003)
	Ideal Employer Judgement	
	Willingness to Approach and Learn More	
Employment Intention towards Private Enterprises (EIPE)	Willingness to Apply	Khan and Muktar (2020) Nguyen Ngoc et al. (2022)
	Willingness to Continue the Job-Seeking Process	
	Willingness to Accept an Offer	
	Priority Choice Tendency	

The target respondents for this study were Chinese fresh graduates and university graduates within two years of graduation. To be eligible, respondents were required to have entered the stage of actively seeking employment, preparing to seek employment, or seriously considering their post-graduation career direction within the past year, and their primary employment target had to be the labour market in mainland China. This criterion was adopted because the present study examines graduates' cognitive evaluations of private-enterprise employment and their intention to seek such employment.

Only individuals who have already entered the actual career decision-making stage and possess a basic capacity to evaluate different types of employers can provide responses that are meaningfully aligned with the research question. The determination of sample size followed the minimum sample size logic commonly applied in PLS-SEM research. On the one hand, according to the “10 times rule”, the minimum sample size should be no less than ten times the largest number of structural paths directed at any single endogenous construct. In the present study, organisational attractiveness is predicted by three exogenous paths, while intention to seek employment in private enterprises is predicted by four paths. Based on the larger of the two, the minimum sample size should therefore not be lower than 40. On the other hand, given that this rule is relatively lenient and that the formal model includes multiple latent constructs and mediating paths, the present study adopted a sample size far exceeding the minimum threshold in order to enhance the credibility of parameter estimation and the stability of model testing. Data for the formal survey were collected between 13 December 2025 and 18 January 2026. The questionnaire was distributed online to eligible members of the target population, and after collection, the responses were screened on the basis of eligibility criteria, completeness of response, and overall answer quality. Only valid samples were retained for the subsequent empirical analysis.

Table 2. Summary of Reliability Analysis (Pilot Test N=45)

Dimension	Item Codes	Number of Items	Cronbach's Alpha
PHC	PHC1–PHC5	5	0.842
PMC	PMC1–PMC5	5	0.861
PES	PES1–PES5	5	0.803
OA	OA1–OA5	5	0.879
EIPE	EIPE1–EIPE5	5	0.891
Full instrument	PHC1–EIPE5	25	0.934

Table 3. Summary of Reliability Analysis (Items Test)

Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Decision
PHC1	0.621	0.812	Retain
PHC2	0.657	0.804	Retain
PHC3	0.603	0.817	Retain
PHC4	0.548	0.829	Retain
PHC5	0.689	0.798	Retain
PMC1	0.644	0.835	Retain
PMC2	0.672	0.829	Retain
PMC3	0.618	0.841	Retain
PMC4	0.701	0.823	Retain
PMC5	0.687	0.826	Retain
PES1	0.533	0.781	Retain
PES2	0.589	0.769	Retain
PES3	0.648	0.755	Retain
PES4	0.617	0.762	Retain
PES5	0.664	0.751	Retain
OA1	0.713	0.847	Retain
OA2	0.738	0.841	Retain
OA3	0.691	0.852	Retain
OA4	0.655	0.860	Retain
OA5	0.724	0.844	Retain

Item	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	Decision
EIPE1	0.746	0.869	Retain
EIPE2	0.781	0.861	Retain
EIPE3	0.719	0.873	Retain
EIPE4	0.688	0.880	Retain
EIPE5	0.754	0.867	Retain

Table 4. KMO and Bartlett's Test for the Pilot Test

Test	Value
KMO Measure of Sampling Adequacy	0.861
Bartlett's Test of Sphericity, Approx. Chi-square	1248.372
df	300
Sig.	< 0.001

Table 5. Exploratory Factor Analysis Results

Item	Factor Loading	Communality	Decision
PHC1	0.731	0.566	Retain
PHC2	0.768	0.603	Retain
PHC3	0.715	0.548	Retain
PHC4	0.681	0.501	Retain
PHC5	0.782	0.618	Retain
PMC1	0.744	0.577	Retain
PMC2	0.771	0.611	Retain
PMC3	0.706	0.539	Retain
PMC4	0.793	0.634	Retain
PMC5	0.759	0.592	Retain
PES1	0.662	0.474	Retain
PES2	0.703	0.521	Retain
PES3	0.748	0.579	Retain
PES4	0.721	0.548	Retain
PES5	0.756	0.587	Retain
OA1	0.801	0.649	Retain
OA2	0.826	0.688	Retain
OA3	0.774	0.619	Retain
OA4	0.741	0.571	Retain
OA5	0.813	0.667	Retain
EIPE1	0.834	0.701	Retain
EIPE2	0.861	0.734	Retain
EIPE3	0.809	0.662	Retain
EIPE4	0.773	0.618	Retain
EIPE5	0.842	0.709	Retain

The pre-test results indicate that the questionnaire developed in this study demonstrates satisfactory measurement quality and structural stability overall, and is therefore capable of capturing the underlying meanings of the latent constructs with reasonable effectiveness. First, in terms of reliability, the Cronbach's alpha coefficients for the five core dimensions were 0.842, 0.861, 0.803, 0.879, and 0.891, respectively, all exceeding the threshold of 0.80. The overall scale achieved a Cronbach's alpha of 0.934, indicating strong internal consistency and suggesting that the questionnaire items exhibited a high degree of coherence and stability. Further item analysis showed that the corrected item–total correlations for all items were above 0.50, and the deletion of any individual item did not produce a meaningful increase in Cronbach's alpha. This suggests that all items were effective in reflecting their respective constructs and that no obviously weak or problematic items required deletion or revision. Accordingly, all items were retained for the formal survey.

With regard to validity, the Kaiser–Meyer–Olkin value was 0.861, and Bartlett's test of sphericity was statistically significant, indicating that the sample data were suitable for factor analysis. In addition, the results of the exploratory factor analysis showed that all items loaded above 0.662 on their respective factors, while all communalities exceeded 0.40. These findings suggest that the measurement items were strongly associated with their corresponding latent factors and were able to explain the structural content of the constructs effectively. Overall, the questionnaire exhibited desirable reliability and initial structural validity at the pre-test stage, indicating that the instrument was well designed, that the wording of the items was sufficiently clear, and that the allocation of items to constructs was appropriately specified. This provides a sound measurement foundation for the subsequent administration of the formal survey and the empirical analysis that follows.

5. Descriptive Statistical Analysis

A total of 446 valid responses were obtained in this study. All respondents met the screening criteria, meaning that they were either fresh graduates or individuals within two years of graduation, had seriously considered employment issues during the past year, and identified the labour market in mainland China as their primary employment target. This indicates a high degree of correspondence between the sample and the target population. In terms of basic demographic characteristics, the gender distribution was relatively balanced, with male respondents accounting for 49.55% and female respondents accounting for 48.65%. With regard to graduation status, respondents graduating in the current year constituted the largest group, representing 50.00% of the sample, followed by those who had graduated within one year at 30.72%, and those who had graduated between one and two years earlier at 19.28%. In terms of educational attainment, bachelor's degree holders formed the majority, accounting for 61.66% of the sample, followed by junior college diploma holders at 19.51%, master's degree holders at 16.59%, and doctoral degree holders at 2.24%. The sample also showed diversity in institutional background. The largest proportion came from ordinary second-tier undergraduate institutions, accounting for 22.42%, followed by higher vocational colleges at 17.26% and ordinary first-tier undergraduate institutions at 16.59%. In terms of disciplinary background, engineering was the most represented field, accounting for 27.58%, followed by science, management, economics and finance, and medicine. Overall, the sample provides reasonably broad coverage of the educational and disciplinary structure commonly found among contemporary university graduates in China.

With regard to respondents' developmental background and employment preparation, most participants came from prefecture-level urban areas, county-level towns, and provincial capitals or sub-provincial cities, while respondents from townships and rural areas also accounted for a certain proportion, suggesting a satisfactory degree of regional diversity within the sample. In terms of preferred long-term employment location, most respondents indicated a preference for new first-tier cities, second-tier cities, and first-tier cities, reflecting a general inclination towards urban platforms offering stronger development opportunities. Regarding family background, annual household income was concentrated primarily in the RMB 100,000 to RMB 200,000 range, followed by households earning below RMB 100,000, indicating that the sample was mainly drawn from lower-middle- to middle-income families. In terms of internship experience, most respondents had undertaken at least one internship, and more than half had previously interned or worked part-time in private enterprises. This suggests that the majority had already acquired some degree of direct exposure to, and practical understanding of, the employment market, particularly the private sector. At the same time, with regard to preferred types of employing organisations, state-owned enterprises, private enterprises, and government agencies or public institutions all accounted for relatively substantial

proportions. This indicates a certain degree of diversity in current graduate employment preferences, although such preferences continue to be shaped in important ways by considerations of stability, development opportunities, and the institutional environment.

Table 6. Demographic Analysis Results (N = 446)

Item	Category	Frequency	Percentage (%)
Gender	Male	221	49.55
	Female	217	48.65
	Other / Prefer not to say	8	1.79
Graduation status	Graduating this year	223	50
	Graduated within 1 year	137	30.72
	Graduated within 1–2 years	86	19.28
Highest educational qualification	Diploma / Junior college	87	19.51
	Bachelor's degree	275	61.66
	Master's degree	74	16.59
	Doctoral degree	10	2.24
Type of graduating institution	Regular second-tier undergraduate university	100	22.42
	Higher vocational / junior college institution	77	17.26
	Regular first-tier undergraduate university	74	16.59
	Other institution types (combined)	195	43.72
Field of study	Engineering	123	27.58
	Science	55	12.33
	Management	52	11.66
	Economics / Finance	36	8.07
	Medicine	34	7.62
	Other fields (combined)	146	32.74
Place of origin	Prefecture-level city urban area	123	27.58
	County town	96	21.52
	Provincial capital / sub-provincial city	75	16.82
	Township	66	14.8
	Rural area	62	13.9
	Municipality directly under the central government	24	5.38
Intended long-term job location	New first-tier city	126	28.25
	Second-tier city	113	25.34
	First-tier city	90	20.18
	Other city tiers / Undecided	117	26.23

Item	Category	Frequency	Percentage (%)
Willingness to work in a different province	Very willing	98	21.97
	Quite willing	130	29.15
	Neutral	106	23.77
	Not very willing / Not willing at all	112	25.11
Family's permanent residence region	Eastern China	171	38.34
	Central China	121	27.13
	Western China	122	27.35
	Northeastern China	32	7.17
Annual household income	Below RMB 100,000	99	22.2
	RMB 100,000–200,000	142	31.84
	Other income levels / Prefer not to say	205	45.96
Family members working in the public sector	Yes	143	32.06
	No	253	56.73
	Not sure	50	11.21
Internship experience	None	74	16.59
	One internship	141	31.61
	Two internships	127	28.48
	Three or more internships	104	23.32
Internship / part-time experience in private enterprises	Yes	236	52.91
	No	210	47.09
Priority employer type in job search	State-owned enterprise	107	23.99
	Private enterprise	95	21.3
	Government agency / public institution	88	19.73
	Undecided	86	19.28
	Other types	70	15.7
Preparation for civil service / public institution / teacher recruitment / SOE examinations	Currently preparing	144	32.29
	Have taken such examinations	116	26.01
	Have not prepared	186	41.7
Perception of the current employment situation	Relatively difficult	181	40.58
	Very difficult	108	24.22
	Other evaluations	157	35.2
Most important factor in job choice	Career development	99	22.2
	Job stability	92	20.63
	Salary and benefits	83	18.61
	Other factors	172	38.56

6. SEM Results

The results for both the direct and indirect effects show that the core paths proposed in this study are statistically supported overall. This indicates that perceived job security, perceived humanistic care, and perceived management compliance not only exert direct effects on fresh graduates' intention to seek employment in private enterprises, but also operate indirectly through the mediating role of organisational attractiveness. More specifically, among the direct effects, perceived job security exerts the strongest influence on intention to seek employment in private enterprises ($\beta = 0.270$, $p = 0.000$), followed by perceived management compliance ($\beta = 0.262$, $p = 0.000$). Organisational attractiveness itself also has a significant positive effect on employment intention ($\beta = 0.208$, $p = 0.000$). Although the coefficient for perceived humanistic care is comparatively smaller, it remains statistically significant ($\beta = 0.118$, $p = 0.013$). With respect to organisational attractiveness, perceived job security again shows the strongest effect ($\beta = 0.338$, $p = 0.000$), followed by perceived humanistic care ($\beta = 0.295$, $p = 0.000$) and perceived management compliance ($\beta = 0.186$, $p = 0.000$). These findings suggest that when graduates evaluate whether a private enterprise is worth approaching and whether it is attractive as a potential employer, their primary concern is whether the organisation can provide a stable, predictable, and low-risk employment environment. At the same time, they also attach considerable importance to whether the firm demonstrates humanistic care and standardised management practices. Turning to the indirect effects, the indirect influences of perceived job security, perceived humanistic care, and perceived management compliance on intention to seek employment in private enterprises through organisational attractiveness are 0.070, 0.061, and 0.039, respectively, and all are statistically significant. This indicates that organisational attractiveness plays a stable mediating role between the three antecedent variables and employment intention, with the mediating effect of perceived job security being the most pronounced. In other words, graduates do not decide whether to choose a private enterprise on the basis of a single condition alone. Rather, they first form an overall evaluation of the organisation's attractiveness, which is then translated into a more concrete employment preference.

The results for explanatory power, effect sizes, and collinearity diagnostics further suggest that the overall model demonstrates satisfactory explanatory capacity and statistical robustness. The R^2 value for employment intention is 0.537, with an adjusted R^2 of 0.533, indicating that the model explains 53.7% of the variance in intention to seek employment in private enterprises. The R^2 value for organisational attractiveness is 0.517, with an adjusted R^2 of 0.513, suggesting that the three antecedent variables also provide substantial explanatory power for organisational attractiveness. Taken together, these results indicate that the model possesses a reasonably strong predictive capability. In terms of effect sizes, most paths fall within the small to moderate range. Among them, perceived job security has the largest effect size on organisational attractiveness ($f^2 = 0.120$), suggesting that this path is not only statistically significant but also substantively important in practical explanatory terms. Perceived management compliance and perceived job security have identical effect sizes on employment intention ($f^2 = 0.071$), while perceived humanistic care shows an effect size of 0.086 on organisational attractiveness, which also indicates meaningful practical relevance. By comparison, the effect size of organisational attractiveness on employment intention is 0.045, that of perceived management compliance on organisational attractiveness is 0.036, and that of perceived humanistic care on employment intention is 0.013. Although these may be classified as relatively small effects, their statistical significance suggests that they remain important components of the overall mechanism. At the same time, the VIF values for all outer-model items range approximately from 1.089 to 1.278, while the VIF values for the inner-model paths range from 1.970 to 2.287. All of these values are well below commonly accepted thresholds, indicating that there are no serious multicollinearity concerns at either the measurement or the structural level. Accordingly, the path estimates among the variables may be regarded as sufficiently independent and reliable. Overall, the empirical results not only support the main theoretical hypotheses, but also demonstrate that when fresh graduates evaluate private enterprises, job security constitutes the most central basis of judgment, while management compliance and humanistic care jointly shape their perceptions of organisational attractiveness and their eventual employment intention through both direct and indirect pathways.

Table 7. Path Coefficient Results

	(O)	(M)	(STDEV)	(O/STDEV)	P
Organisational Attractiveness -> Employment Intention towards_Private Enterprises	0.208	0.207	0.051	4.065	0.000
Perceived Employment Security -> Employment Intention towards_Private Enterprises	0.270	0.270	0.050	5.372	0.000
Perceived Employment Security -> Organisational Attractiveness	0.338	0.339	0.044	7.643	0.000
Perceived Humanistic Care -> Employment Intention towards_Private Enterprises	0.118	0.121	0.047	2.496	0.013
Perceived Humanistic Care -> Organisational Attractiveness	0.295	0.297	0.048	6.098	0.000
Perceived Management Compliance -> Employment Intention towards_Private Enterprises	0.262	0.262	0.047	5.620	0.000
Perceived Management Compliance -> Organisational Attractiveness	0.186	0.186	0.047	3.984	0.000

Table 8. Path Coefficient Results (Indirect Path)

	(O)	(M)	(STDEV)	(O/STDEV)	P
Perceived Employment Security -> Employment Intention towards_Private Enterprises	0.070	0.070	0.019	3.610	0.000
Perceived Humanistic Care -> Employment Intention towards_Private Enterprises	0.061	0.061	0.018	3.403	0.001
Perceived Management Compliance -> Employment Intention towards_Private Enterprises	0.039	0.039	0.015	2.655	0.008

Table 9. Summary of R^2 , Adjusted R^2 , and f^2 Results

Indicator	Path / Construct	O	M	STDEV	t	p
R^2	Employment Intention towards Private Enterprises	0.537	0.545	0.03	18.161	0.000
	Organisational Attractiveness	0.517	0.523	0.032	16.352	0.000
Adjusted R^2	Employment Intention towards Private Enterprises	0.533	0.54	0.03	17.857	0.000
	Organisational Attractiveness	0.513	0.52	0.032	16.139	0.000
f^2	Organisational Attractiveness → Employment Intention towards Private Enterprises	0.045	0.048	0.023	1.967	0.007
	Perceived Employment Security → Employment Intention towards Private Enterprises	0.071	0.075	0.029	2.474	0.013
	Perceived Employment Security → Organisational Attractiveness	0.12	0.124	0.034	3.485	0.000
	Perceived Humanistic Care → Employment Intention towards Private Enterprises	0.013	0.016	0.011	1.153	0.009
	Perceived Humanistic Care → Organisational Attractiveness	0.086	0.09	0.031	2.781	0.005
	Perceived Management Compliance → Employment Intention towards Private Enterprises	0.071	0.075	0.027	2.613	0.009
	Perceived Management Compliance → Organisational Attractiveness	0.036	0.039	0.019	1.867	0.006

Table 10. Summary of Outer and Inner Model VIF Results

Model	Indicator / Path	O	M	2.50%	97.50%
Outer	EIPE1	1.105	1.115	1.059	1.184
	EIPE2	1.149	1.159	1.09	1.244
	EIPE3	1.089	1.1	1.047	1.167
	EIPE4	1.161	1.172	1.099	1.264
	EIPE5	1.2	1.211	1.127	1.314
	OA1	1.155	1.167	1.092	1.264
	OA2	1.256	1.268	1.174	1.379
	OA3	1.225	1.237	1.146	1.345
	OA4	1.19	1.201	1.12	1.3
	OA5	1.265	1.277	1.183	1.391
	PES1	1.143	1.155	1.086	1.237
	PES2	1.117	1.127	1.066	1.205
	PES3	1.258	1.27	1.177	1.383
	PES4	1.278	1.291	1.192	1.405
	PES5	1.246	1.259	1.167	1.372
	PHC1	1.259	1.271	1.178	1.379
	PHC2	1.15	1.161	1.095	1.244
	PHC3	1.14	1.15	1.087	1.228
	PHC4	1.162	1.175	1.099	1.269
	PHC5	1.105	1.116	1.062	1.182
	PMC1	1.154	1.167	1.094	1.257
	PMC2	1.181	1.194	1.115	1.29
	PMC3	1.143	1.154	1.085	1.24
	PMC4	1.185	1.196	1.119	1.287
	PMC5	1.129	1.14	1.073	1.222
Inner	OA → EIPE	2.069	2.107	1.846	2.402
	PES → EIPE	2.206	2.239	1.934	2.577
	PES → OA	1.97	1.991	1.748	2.261
	PHC → EIPE	2.287	2.316	2.048	2.629
	PHC → OA	2.107	2.126	1.876	2.41
	PMC → EIPE	2.069	2.093	1.834	2.381
	PMC → OA	1.997	2.016	1.769	2.293

Graphic 2. SEM output

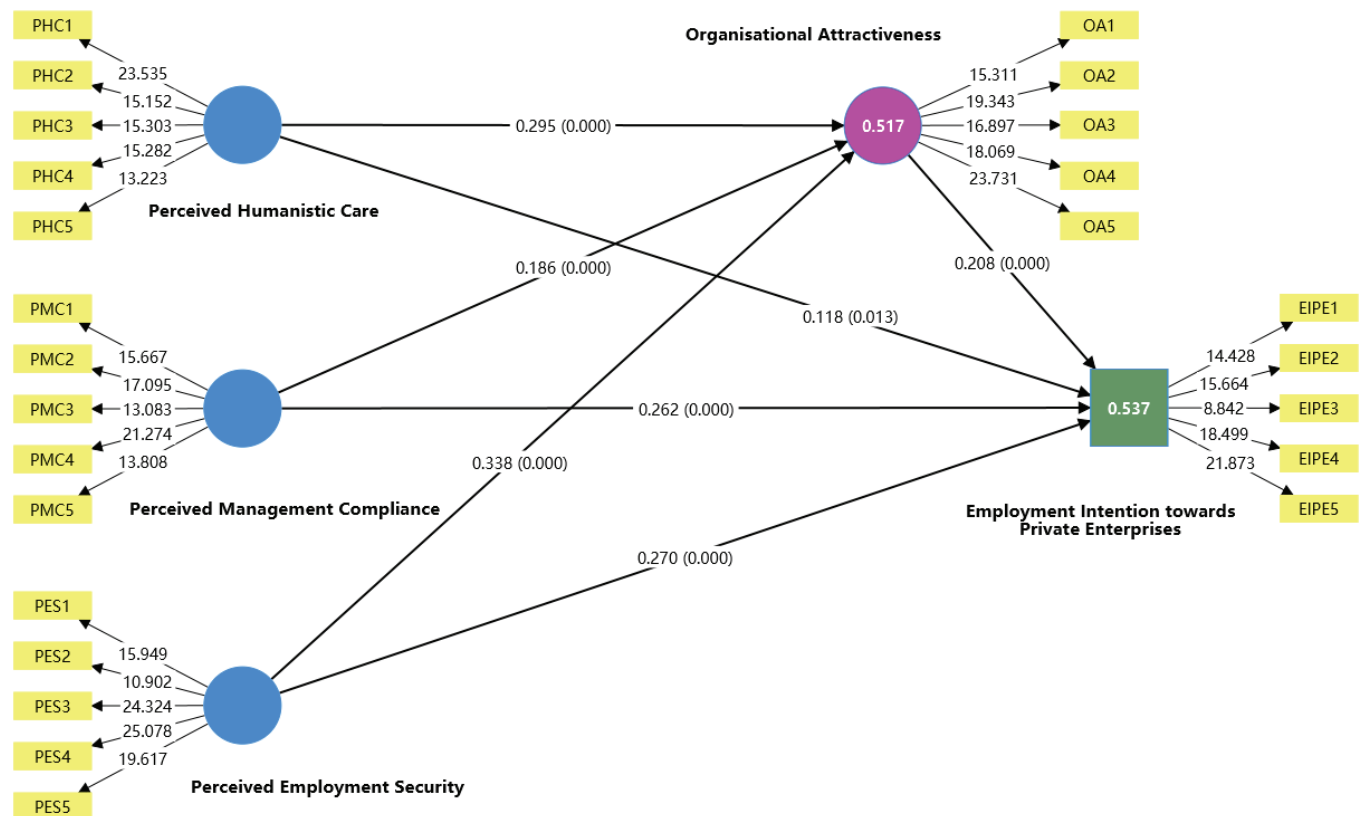


Table 11. Hypothesis Testing Results

Hypothesis	Path	Beta	t-value	p-value	Decision
H1	PHC → EIPE	0.118	2.496	0.013	Supported
H2	PMC → EIPE	0.262	5.62	0.000	Supported
H3	PES → EIPE	0.27	5.372	0.000	Supported
H4	PHC → OA	0.295	6.098	0.000	Supported
H5	PMC → OA	0.186	3.984	0.000	Supported
H6	PES → OA	0.338	7.643	0.000	Supported
H7	OA → EIPE	0.208	4.065	0.000	Supported
H8a	PHC → OA → EIPE	0.061	3.403	0.001	Supported
H8b	PMC → OA → EIPE	0.039	2.655	0.008	Supported
H8c	PES → OA → EIPE	0.07	3.61	0.000	Supported

According to the empirical findings, all hypotheses proposed in this study are supported. Perceived humanistic care, perceived management compliance, and perceived job security all exert significant positive effects on Chinese fresh graduates' intention to seek employment in private enterprises. Among these factors, perceived job security shows the strongest direct effect, followed by perceived management compliance, while perceived humanistic care has a relatively weaker, though still statistically significant, influence. This suggests that in the context of seeking employment in private enterprises, fresh graduates are concerned not only with whether an organisation demonstrates human warmth and consideration, but also with whether its management practices are well regulated and whether the employment environment is stable and predictable. At

the same time, perceived humanistic care, perceived management compliance, and perceived job security all significantly enhance organisational attractiveness, with perceived job security again exerting the strongest effect. This indicates that when graduates form an overall attractiveness evaluation of a private enterprise, the most important consideration remains whether the organisation can offer a reassuring and stable starting point for their careers. Furthermore, organisational attractiveness has a significant positive effect on intention to seek employment in private enterprises and serves as a significant mediator in the relationships between perceived humanistic care, perceived management compliance, perceived job security, and intention to seek employment in private enterprises. These findings suggest that graduates do not form employment intentions immediately upon perceiving a single favourable organisational attribute. Rather, they first translate such positive perceptions into an overall judgment of organisational attractiveness, which then develops into a more definite intention to pursue employment in private enterprises.

7. Conclusion and Implications

The reason why fresh graduates today generally display a stronger preference for public-sector positions while remaining cautious about, or even avoiding, private enterprises cannot be reduced simply to a generic “preference for stability”. More fundamentally, within their real-world employment perceptions, public-sector organisations are typically associated with stronger expectations of security, clearer rules and institutional order, and lower career risk. For graduates who are only beginning to enter the labour market, occupational capital remains limited and tolerance for uncertainty is relatively low. As a result, they are inclined to prioritise organisational settings characterised by clear boundaries, formalised systems, stronger protection, and a more predictable future trajectory. By contrast, the insufficient attractiveness of many private enterprises does not stem solely from salary levels. Rather, graduates commonly worry that such firms may offer inadequate employment security, weak managerial standardisation, and an organisational environment marked by excessive arbitrariness. They may also anticipate unstable positions, ambiguous performance criteria, opaque promotion mechanisms, normalised overtime, insufficient protection of rights and interests, and even a lack of basic respect after entry. In other words, what private enterprises truly lack when competing with the public sector is neither recruitment rhetoric nor the superficial packaging of benefits, but a form of security, order, and trust that graduates can clearly perceive. The findings of this study confirm precisely this point: perceived job security is the most critical factor shaping graduates' intention to seek employment in private enterprises, followed by perceived management compliance and then perceived humanistic care, and all three further influence final choice through organisational attractiveness. This suggests that graduates are not inherently resistant to private enterprises. Rather, under prevailing labour market perceptions, many private firms have not yet developed sufficiently strong organisational competitiveness to offset the institutional advantages associated with the public sector.

Accordingly, if private enterprises wish to establish genuine competitiveness under the current pattern in which graduates cluster towards the public sector, they must shift from merely “offering jobs” to “offering a trustworthy career starting point”. First, employment security must be translated from an abstract promise into concrete institutional arrangements. This includes clearly defined labour contracts, reduced arbitrariness in job reassignment and dismissal, timely and full payment of wages and statutory social insurance contributions, as well as transparent career development paths and relatively stable expectations regarding job continuity. Graduates need to feel that entering a private enterprise is not a high-risk trial-and-error decision. Second, private firms must improve the level of management compliance by establishing open and transparent recruitment procedures, clear and consistent performance evaluation systems, lawful and standardised employment practices, and predictable reward and sanction mechanisms. Such measures are essential to dispel graduates' negative impressions that private firms are characterised by arbitrary management, blurred rules, and governance driven by personal discretion rather than institutional procedures. Third, private enterprises must address a long-neglected humanistic deficit by demonstrating respect for young employees throughout recruitment, onboarding, mentoring, communication, and developmental support. Fresh graduates need to feel that they are not replaceable low-cost labour, but formal organisational members worthy of investment and cultivation. Ultimately, if private enterprises intend to attract talent away from the public sector, they cannot rely solely on competing in terms of pay, pace, or slogans. They must instead build, at the organisational level, a comprehensive form of attractiveness capable of standing against the appeal of public-sector institutions. Graduates must

genuinely believe that such jobs are not only available, but also stable, developmental, and fair. Only when private enterprises establish clear, consistent, and perceptible competitive advantages in security, standardisation, and humanistic care will the balance of graduate job choice begin to shift from risk-avoidant entry into the public sector towards an active preference for private-enterprise employment.

8. Research Limitations

This study is subject to several limitations. First, in terms of research design, the study relies on a one-shot cross-sectional survey and self-reported data to measure perceived humanistic care, perceived management compliance, perceived job security, organisational attractiveness, and intention to seek employment in private enterprises. Although the pre-test, reliability and validity assessments, and structural model results all indicate satisfactory scale quality and model robustness, such a design is more suitable for identifying statistical associations among variables than for making stronger causal inferences. It is also limited in its ability to capture the dynamic adjustments in fresh graduates' job choice judgments over time, across different employment stages, and under changing external conditions. Second, with regard to sample scope, although the study includes 446 valid responses and covers graduates with different educational levels, institutional backgrounds, disciplinary fields, places of origin, and family conditions, the sample was still collected primarily through an online questionnaire and largely consisted of fresh graduates and individuals within two years of graduation who had already entered the job-seeking stage and were mainly oriented towards the labour market in mainland China. Accordingly, caution is still required when extending the findings to broader youth populations, different regional labour markets, private enterprises in different industries, or young people who are not university graduates.

Third, in terms of variable specification, this study focuses on three forms of organisational perception, namely perceived job security, perceived management compliance, and perceived humanistic care, and explains their effects on intention to seek employment in private enterprises through organisational attractiveness. Although this framework responds well to the current reality of public-sector preference and relatively weak intention towards private-enterprise employment among graduates, it does not incorporate into the same model a number of other potentially important factors, such as salary and benefits, employer branding, family expectations, social comparison, opportunities for urban household registration, macroeconomic expectations, and individual risk preference. As a result, the explanation offered for private-enterprise employment intention should not be regarded as exhaustive. Finally, this study addresses graduates' intention to seek employment in private enterprises rather than actual entry behaviour or post-entry retention outcomes. The findings therefore reveal primarily the psychological mechanisms operating at the pre-employment decision stage, rather than the full behavioural chain from job search to offer acceptance and subsequent retention. Future research may build on this study by employing longitudinal tracking, experimental designs, multi-source data, or comparative analyses across different types of enterprises so as to enhance external validity and deepen explanatory insight.

Funding

No

Conflict of Interests

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

Reference

- [1] Dash, G., & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, 173, 121092. <https://doi.org/10.1016/j.techfore.2021.121092>
- [2] Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18(4), 694–734. <https://doi.org/10.2307/258595>
- [3] Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *An introduction to structural equation modeling*. Springer. https://doi.org/10.1007/978-3-030-80519-7_1
- [4] Highhouse, S., Lievens, F., & Sinar, E. F. (2003). Measuring attraction to organizations. *Educational and Psychological Measurement*, 63(6), 986–1001. <https://doi.org/10.1177/0013164403258403>

- [5] Jones, D. A., Newman, A., Shao, R., & Cooke, F. L. (2016). The applications of person–organization fit to the study of job seekers and applicants: A review and future research agenda. *Frontiers in Psychology*, 7, Article 426. <https://doi.org/10.3389/fpsyg.2016.00426>
- [6] Khan, M. H., & Muktar, S. N. (2020). Mediating role of organizational attractiveness on the relationship between green recruitment and job pursuit intention among students of Universiti Teknologi Malaysia. *Cogent Business & Management*, 7(1), 1832811. <https://doi.org/10.1080/23311975.2020.1832811>
- [7] Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610. <https://doi.org/10.1177/001316447003000308>
- [8] Kurtessis, J. N., Eisenberger, R., Ford, M. T., Buffardi, L. C., Stewart, K. A., & Adis, C. S. (2017). Perceived organizational support: A meta-analytic evaluation of organizational support theory. *Journal of Management*, 43(6), 1854–1884. <https://doi.org/10.1177/0149206315575554>
- [9] Nguyen Ngoc, T., Dung, M. V., Rowley, C., & Pejić Bach, M. (2022). Generation Z job seekers' expectations and their job pursuit intention: Evidence from transition and emerging economy. *International Journal of Engineering Business Management*, 14, 1–13. <https://doi.org/10.1177/18479790221112548>
- [10] Roehling, M. V., Cavanaugh, M. A., Moynihan, L. M., & Boswell, W. R. (2000). Job security rights: The effects of specific policies and practices on job seekers' evaluations of employers. *Journal of Organizational Behavior*, 21(7), 777–800. [https://doi.org/10.1002/1099-1379\(200011\)21:7](https://doi.org/10.1002/1099-1379(200011)21:7)
- [11] Ruel, E., Wagner, W. E., & Gillespie, B. J. (2019). How can pretesting and pilot testing improve reliability and validity of your survey? In *The practice of survey research* (2nd ed.). SAGE. <https://doi.org/10.4135/9781506348803.n59>
- [12] Thomas, K. M., & Lucas, K. (2019). Development and validation of the Workplace Dignity Scale. *Group & Organization Management*, 44(1), 72–111. <https://doi.org/10.1177/1059601118807784>
- [13] Truxillo, D. M., Bauer, T. N., Campion, M. A., & Paronto, M. E. (2004). The importance of organizational justice in personnel selection: Defining when selection fairness really matters. *International Journal of Selection and Assessment*, 12(1–2), 39–53. <https://doi.org/10.1111/j.0965-075X.2004.00262.x>
- [14] Vogel, R., & Satzger, M. (2024). What drives the attractiveness of public and private employers? Comparative evidence from an online employer review platform. *The American Review of Public Administration*, 54(2), 95–112. <https://doi.org/10.1177/02750740231206805>
- [15] Willis, G. B. (2011). Cognitive interviewing to test and refine questionnaires. In J. A. Madans, K. Miller, A. Maitland, & G. Willis (Eds.), *Question evaluation methods: Contributing to the science of data quality* (pp. 3–22). Wiley. <https://doi.org/10.1111/j.1525-1446.2010.00938.x>
- [16] Ahamad, F., Saini, G. K., & Jawahar, I. M. (2023). Interactive influence of work–life balance benefits, employee recommendation, and job attributes on employer attractiveness and job pursuit intentions: Two experiments. *Asian Business & Management*. <https://doi.org/10.1057/s41291-022-00184-4>
- [17] Bustaman, H. A., Teng, S.-S., & Abdullah, M. I. (2020). Job seeker attraction to organizational justice mediated by organizational reputation. *Cogent Psychology*, 7(1), 1816255. <https://doi.org/10.1080/23311908.2020.1816255>
- [18] Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18(4), 694–734. <https://doi.org/10.2307/258595>
- [19] Jain, N., & Bhatt, P. (2015). Employment preferences of job applicants: Unfolding employer branding determinants. *Journal of Management Development*, 34(6), 634–652. <https://doi.org/10.1108/JMD-09-2013-0106>
- [20] Jones, D. A., Newman, A., Shao, R., & Cooke, F. L. (2016). The applications of person–organization fit to the study of job seekers and applicants: A review and future research agenda. *Frontiers in Psychology*, 7, Article 426. <https://doi.org/10.3389/fpsyg.2016.00426>
- [21] Jones, D. A., Willness, C. R., & Madey, S. (2014). Why are job seekers attracted by corporate social performance? Experimental and field tests of three signal-based mechanisms. *Academy of Management Journal*, 57(2), 383–404. <https://doi.org/10.5465/amj.2011.0848>

- [22] Khan, M. H., & Muktar, S. N. (2020). Mediating role of organizational attractiveness on the relationship between green recruitment and job pursuit intention among students of Universiti Teknologi Malaysia. *Cogent Business & Management*, 7(1), 1832811. <https://doi.org/10.1080/23311975.2020.1832811>
- [23] Kumari, S., & Saini, G. K. (2018). Do instrumental and symbolic factors interact in influencing employer attractiveness and job pursuit intention? *Career Development International*, 23(4), 444–462. <https://doi.org/10.1108/CDI-03-2018-0069>
- [24] Mauger, Y., & Bryant-Lees, K. (2022). What are job seekers attracted to? Examining the importance of organizational attributes across four types of job seekers in Poland and France. *Baltic Journal of Management*, 17(3), 338–355. <https://doi.org/10.1108/BJM-05-2021-0193>
- [25] Nguyen Ngoc, T., Dung, M. V., Rowley, C., & Pejić Bach, M. (2022). Generation Z job seekers' expectations and their job pursuit intention: Evidence from transition and emerging economy. *International Journal of Engineering Business Management*, 14, 1–13. <https://doi.org/10.1177/18479790221112548>
- [26] Ministry of Education of the People's Republic of China. (2024, November 12). Notice on employment and entrepreneurship work for the 2025 cohort of college graduates. https://www.moe.gov.cn/srcsite/A15/s3265/202411/t20241112_1162526.html
- [27] Ministry of Human Resources and Social Security of the People's Republic of China, Ministry of Education of the People's Republic of China, & Ministry of Finance of the People's Republic of China. (2025, May 7). Notice on further improving employment and entrepreneurship work for college graduates and other young people in 2025. <https://chinajob.mohrss.gov.cn/h5/c/2025-05-07/432972.shtml>
- [28] All-China Federation of Industry and Commerce. (2025, April 9). 2025 annual launch meeting of the “Hundreds of Cities, Thousands of Schools, and Tens of Thousands of Enterprises” employment promotion campaign and Hubei special recruitment event. https://www.acfic.org.cn/ztzlh/2023mqgxcjy/2023mqgxcjy_dt/202504/t20250409_314973.html
- [29] Beijing Municipal Human Resources and Social Security Bureau. (2026, January 12). Promoting fuller employment for young people through high-quality development. https://tyjrswj.beijing.gov.cn/fwzl/jycy/202601/t20260112_4422284.html
- [30] Ministry of Education of the People's Republic of China. (2024a, November 14). National conference on employment and entrepreneurship work for the 2025 cohort of college graduates convened. https://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/moe_1485/202411/t20241114_1162975.html
- [31] Ministry of Education of the People's Republic of China. (2024b, November 15). The number of college graduates in the 2025 cohort is expected to reach 12.22 million. https://www.moe.gov.cn/jyb_xwfb/s5147/202411/t20241115_1163118.html
- [32] National Development and Reform Commission. (2025, November 18). A set of data shows the simultaneous rise in the scale and quality of the private economy. https://www.ndrc.gov.cn/wsdwhfz/202511/t20251118_1401698.html

Appendix: Questionnaire Survey

(Data collected from 12/2025 - 01/2026 only)

SURVEY QUESTIONNAIRE

Chinese Fresh Graduates' Employment Intentions toward Private Enterprises

Dear Respondent,

This questionnaire is designed for academic research. It aims to understand how Chinese fresh graduates perceive employment in private enterprises and how these perceptions shape their employment intentions.

The findings will help identify key factors influencing graduates' willingness to work in private enterprises and may provide useful references for improving employer attractiveness, strengthening talent recruitment, and supporting higher-quality graduate employment.

The questionnaire is anonymous. All responses will be used for statistical analysis only and will remain strictly confidential. There are no right or wrong answers. Please answer according to your genuine views and experiences.

Estimated completion time: 5-8 minutes. Thank you very much for your support.

Part I. Screening Questions

Please answer these questions first. Only eligible respondents should continue.

S1. Are you currently a fresh graduate, or have you graduated within the past two years?

- ☐ Yes
- ☐ No

S2. Are you currently seeking employment, preparing to seek employment, or have you seriously considered your post-graduation employment options within the past year?

- ☐ Yes
- ☐ No

S3. How familiar are you with private enterprises as a type of employer?

- ☐ Very familiar
- ☐ Quite familiar
- ☐ Basically familiar
- ☐ Not very familiar
- ☐ Not familiar at all

S4. Is your main employment target the mainland China labour market?

- ☐ Yes
- ☐ No

Part II. Demographic Information

Please tick the option that best describes your situation.

D1. Gender

- ☐ Male
- ☐ Female
- ☐ Other
- ☐ Prefer not to say

D2. Year of birth

- ☐ 1998 or earlier
- ☐ 1999
- ☐ 2000
- ☐ 2001
- ☐ 2002
- ☐ 2003
- ☐ 2004 or later

D3. Current graduation status

- ☐ Graduating this year
- ☐ Graduated within 1 year
- ☐ Graduated within 1-2 years

D4. Marital status

- ☐ Single
- ☐ Married
- ☐ Other

D5. Highest educational qualification

- ☐ Diploma / Junior College
- ☐ Bachelor's degree
- ☐ Master's degree
- ☐ Doctoral degree

D6. Which category best describes your graduating institution?

- ☐ Double First-Class university
- ☐ Former 985 university
- ☐ Former 211 university
- ☐ Regular first-tier undergraduate university
- ☐ Regular second-tier undergraduate university
- ☐ Higher vocational / junior college institution
- ☐ Private undergraduate institution / independent college
- ☐ Overseas university
- ☐ Other

D7. Field of study

- ☐ Management
- ☐ Economics / Finance

- ☐ Law
- ☐ Literature / Journalism / Communication
- ☐ Education
- ☐ Science
- ☐ Engineering
- ☐ Medicine
- ☐ Agriculture
- ☐ Arts
- ☐ Interdisciplinary / Emerging field
- ☐ Other

D8. How well does your major match market-oriented enterprise positions?

- ☐ Very well
- ☐ Quite well
- ☐ Moderately
- ☐ Poorly
- ☐ Very poorly

D9. Your approximate academic ranking during your studies

- ☐ Top 10%
- ☐ Top 10%-30%
- ☐ Top 30%-50%
- ☐ Bottom 50%
- ☐ Prefer not to say

D10. Did you hold any student leadership roles (e.g., class, student union, club, or association positions)?

- ☐ Yes
- ☐ No

D11. Your place of origin is best described as

- ☐ Municipality directly under the central government
- ☐ Provincial capital / sub-provincial city
- ☐ Prefecture-level city urban area
- ☐ County town
- ☐ Township
- ☐ Rural area

D12. Your intended long-term job location is

- ☐ First-tier city
- ☐ New first-tier city
- ☐ Second-tier city
- ☐ Third-tier city
- ☐ Fourth-tier or below city
- ☐ Not yet decided

D13. Willingness to work in a different province

- ☐ Very willing
- ☐ Quite willing
- ☐ Neutral
- ☐ Not very willing
- ☐ Not willing at all

D14. Which region does your family's permanent residence belong to?

- ☐ Eastern China
- ☐ Central China
- ☐ Western China
- ☐ Northeastern China

D15. Father's highest educational level

- ☐ Lower secondary or below
- ☐ Upper secondary / technical secondary
- ☐ Junior college
- ☐ Bachelor's degree
- ☐ Master's degree or above
- ☐ Do not know / prefer not to say

D16. Mother's highest educational level

- ☐ Lower secondary or below
- ☐ Upper secondary / technical secondary
- ☐ Junior college
- ☐ Bachelor's degree
- ☐ Master's degree or above
- ☐ Do not know / prefer not to say

D17. Approximate annual household income

- ☐ Below RMB 100,000
- ☐ RMB 100,000-200,000
- ☐ RMB 200,000-300,000
- ☐ RMB 300,000-500,000
- ☐ Above RMB 500,000
- ☐ Prefer not to say

D18. Does your family have members working in government agencies, public institutions, or other public-sector organisations?

- ☐ Yes
- ☐ No
- ☐ Not sure

D19. Internship experience

- ☐ None

- ☐ One internship
- ☐ Two internships
- ☐ Three or more internships

D20. Have you ever interned or worked part-time in a private enterprise?

- ☐ Yes
- ☐ No

D21. Which type of employer do you prioritise most in your job search?

- ☐ Government agency / public institution
- ☐ State-owned enterprise
- ☐ Private enterprise
- ☐ Foreign-funded enterprise
- ☐ Self-employment / entrepreneurship
- ☐ Undecided

D22. Have you prepared for or taken civil service, public institution, teacher recruitment, or state-owned enterprise examinations?

- ☐ Yes, currently preparing
- ☐ Yes, I have taken such examinations
- ☐ No

D23. How do you perceive the current employment situation?

- ☐ Very favourable
- ☐ Relatively favourable
- ☐ Average
- ☐ Relatively difficult
- ☐ Very difficult

D24. Which factor matters most to you when choosing a job?

- ☐ Salary and benefits
- ☐ Job stability
- ☐ Career development
- ☐ Managerial standardisation
- ☐ Work atmosphere
- ☐ City platform
- ☐ Match with major
- ☐ Other

Part III. Main Questionnaire

Please indicate the extent to which you agree with each statement about employment in private enterprises.

Response scale: 1 = Strongly disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly agree

A. Perceived Humane Care (PHC)

No.	Statement	1	2	3	4	5
1	I feel that many private enterprises respect the feelings of young employees in their day-to-day management.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	If employees encounter personal or work-related difficulties, I think private enterprises are generally willing to show understanding.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	I believe private enterprises often treat fresh graduates with a certain degree of empathy and human concern.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	When evaluating a private enterprise, I care about whether it genuinely cares about employees rather than only about performance.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	If a private enterprise makes me feel respected, I would be more willing to work there.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B. Perceived Managerial Compliance (PMC)

No.	Statement	1	2	3	4	5
6	When job hunting, I care greatly about whether a private enterprise complies with labour regulations.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	I pay attention to whether a private enterprise has problems such as excessive overtime, arbitrary penalties, or unfulfilled verbal promises.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	If a private enterprise has clear management procedures and rules, I feel more secure about it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	Whether recruitment procedures, labour contracts, and salary payments are handled properly influences my evaluation of a private enterprise.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	I would be more willing to join private enterprises that are transparent and rule-based in management.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C. Perceived Employment Security (PES)

No.	Statement	1	2	3	4	5
11	I worry that jobs in some private enterprises may not be sufficiently stable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	I pay attention to whether a private enterprise is likely to engage in layoffs, post reductions, or sudden staffing changes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13	If a private enterprise develops in a stable manner, I would be more willing to consider it.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14	I hope that income, job arrangements, and labour relations after joining an employer will be predictable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15	Whether a private enterprise can provide graduates with a reassuring employment experience matters to me.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

D. Organisational Attractiveness (OA)

No.	Statement	1	2	3	4	5
16	If a private enterprise leaves me with a good overall impression, I am willing to learn more about it proactively.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17	I believe some private enterprises are genuinely worthwhile employment options for fresh graduates.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18	When I encounter a private enterprise that I like, I am willing to submit my application.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19	If conditions are suitable, I would recommend such a private enterprise to my classmates or friends.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20	My overall impression of a private enterprise influences whether I include it on my shortlist of job options.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

E. Employment Intention toward Private Enterprises (EIPE)

No.	Statement	1	2	3	4	5
21	During my job search after graduation, I am willing to treat private enterprises as a formal employment option.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22	If job conditions are appropriate, I am willing to give priority to working in a private enterprise.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23	Rather than waiting indefinitely for other opportunities, I am willing to try a suitable private enterprise.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
24	I can accept a private enterprise as my first full-time job after graduation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
25	In future job searching, I hold an open attitude toward working in private enterprises.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Part IV. Supplementary Questions

These items may be used for descriptive or control-variable analysis.

C1. How would you evaluate your overall impression of private enterprises?

- ☐ Very poor
- ☐ Poor
- ☐ Average
- ☐ Good
- ☐ Very good

C2. What is the general attitude of your family members or teachers toward employment in private enterprises?

- ☐ Strongly unsupportive
- ☐ Unsupportive
- ☐ Neutral
- ☐ Supportive
- ☐ Strongly supportive

C3. Compared with state-owned enterprises or public-sector jobs, private enterprises can offer better opportunities for growth.

- ☐ Strongly disagree
- ☐ Disagree
- ☐ Neutral
- ☐ Agree
- ☐ Strongly agree

C4. When seeking employment, do you hesitate to join a private enterprise because of social evaluation or external opinions?

- ☐ Never
- ☐ Rarely
- ☐ Sometimes
- ☐ Often
- ☐ Always

Part V. Open-ended Questions

O1. If you are not very willing to work in a private enterprise, what is your main concern?

O2. If a private enterprise wishes to attract you, what do you most hope it would improve?

Thank you for your participation.

Research on the Evaluation of ESG Practice Effectiveness of Shanxi J Company

Shijie Sun*

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

*Corresponding author: Shijie Sun, 1405935853@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: Shanxi, as a major coal province, has been actively exploring the path of resource transformation and energy development. With the introduction of ESG concepts and increasing attention from listed companies in various countries and regions, Shanxi J Company began disclosing ESG information in 2024. Since the release of its second ESG report in 2025, it has further improved its sustainable disclosure work. This article uses the entropy mutation series method to study the ESG practice effectiveness of Shanxi J Company. The study found that Shanxi J Company has great potential in ESG and should continue to strengthen sustainable information disclosure in the future. According to relevant work requirements, the company's ESG practice level should be effectively improved.

Keywords: Shanxi J Company; ESG; Entropy Value-mutation Series Method

Published: Mar 29, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1174>

1. Introduction

1.1 research background

Energy, as the key material cornerstone of the national economy, its future development will depend on the control of energy. In 2004, the United Nations Global Compact put forward the concept of ESG, which emphasizes the combination of environment, society and corporate governance, and is committed to promoting the sustainable development of enterprises. Today, after 40 years of development, the concept of sustainable development and ESG has increasingly become a topic of concern for enterprises. Shanxi, as a province rich in energy and resources, has always shouldered important responsibilities. In July 2025, he came here again to inspect the work and stressed the following point of view: "building a national energy conversion and comprehensive cooperation reform experimental area has been given an important strategic mission by the central government to Shanxi Province. We need to continue to deepen our understanding, adhere to our determination, and promote the process of transformation and development in an orderly manner." in the process of energy transformation and development, ESG concept will bring more opportunities for enterprises in Shanxi energy industry, and help Shanxi build a new energy system and realize transformation and development.

1.2 research significance

At present, Shanxi is in the critical period of the transformation and development of resource-based economy and energy. How to achieve sustainable development has become an increasingly important topic, and the ESG concept theoretically provides the possibility for the transformation and development of Shanxi. The research of this paper has both theoretical

and practical significance: on the one hand, it is based on the theoretical construction of the evaluation of the effect of ESG practice of Shanxi J company, which lays the foundation for the in-depth study of its ESG performance; On the other hand, with the transformation of the coal industry and the practice of ESG, the coking industry also keeps pace, and the relevant theoretical and practical research should also be carried out simultaneously. It is necessary to combine theory with practice to better grasp how the ESG concept can promote the transformation and upgrading of Shanxi J Company.

1.3 literature review

At present, the research on ESG at home and abroad is very rich, which also reflects the continuous improvement of enterprises' attention to sustainable development. The research on ESG performance, enterprise value and impact path is very rich, and the research is mainly carried out by case study method, entropy mutation progression and other comprehensive evaluation methods, reflecting the academic concern about the impact of enterprise ESG performance on enterprise value.

1.3.1 domestic literature review

The concept of ESG started late in China, so there are few initial studies. In recent years, with the new development concept, China's emphasis on environment and ecology and the "double carbon" goal, the topic of ESG has become a hot topic of research. Starting from the concept of ESG, it has become the "door" to open the research. Liu Xiao and other scholars' research shows that domestic scholars' research on ESG is divided into three categories: ESG investment and enterprise value, ESG information disclosure and social responsibility, green innovation and environmental governance^[1]. And the research system with these three categories as the core has been built in China. Closely related to ESG is CSR, that is, social responsibility. Realizing the collaborative preparation and value maximization of CSR report and ESG report has become an important topic in the current management practice of state-owned enterprises^[2]. The main reference for ESG research is also from the report. In addition, many domestic Master of accounting theses also focus on enterprise ESG performance. The research content covers the relationship between enterprise ESG performance and enterprise value, influence path and enterprise financial disclosure data. Comprehensive evaluation methods are often selected, such as AHP, grey relational analysis, fuzzy evaluation method, expert scoring method, etc^[3]. These methods are also frequently used in modern management.

Generally speaking, domestic scholars' research on ESG covers both qualitative and conceptual theoretical exploration, as well as quantitative analysis taking enterprises as cases. There are also research on ESG performance in combination with hot topics such as new quality productivity and digital economy. The main driving force for domestic enterprises to practice ESG is the regulatory compliance requirements. Although the number of ESG reports disclosed by Chinese enterprises has increased significantly, the quality of disclosure needs to be improved^[4]. The research related to ESG also needs to further improve the quality, link the concept of ESG with all aspects of economic life, form an organic whole, so as to grasp its connotation more comprehensively. Relevant research has also paid attention to the relationship between regional economy and ESG.

1.3.2 foreign literature review

The concept of ESG originated abroad, so the time and maturity of relevant research abroad will be higher than that in China. ESG related concepts have emerged since the 1970s, that is, corporate social responsibility (CSR). In 2004, the United Nations Global Compact officially put forward the ESG concept and formed a series of international industry sustainable information disclosure standards. In 2010, the global reporting initiative advocated enterprises to integrate and disclose financial information and ESG information. Many international institutions such as Dow Jones, Thomson Reuters and FTSE Russell have put forward their own evaluation systems^[5]. With this trend, scholars began to study related issues in depth. Early scholars mainly focused on the study of environmental responsibility, and they only analyzed it from the perspective of compliance with laws and regulations and self-discipline. Bansal and Roth put forward a view that enterprises will make corresponding changes in order to reduce the damage to the ecological environment^[6]. With the in-depth study, scholars have combined corporate environmental responsibility with stakeholder theory and 3P principle. Typical examples of this are Henriques and sadorsky, who proposed that environmental responsibility is a process in which enterprises have a positive impact on society through open and transparent relationship management with stakeholders^[7]. With the development of the

times, the theory is also constantly enriched. The question of traditional neoclassical environmental economics forms the “Porter Hypothesis Theory”, which believes that appropriate environmental regulation can encourage enterprises to make more innovation attempts. With the continuous deepening of the concept of green and sustainable development, the behavior of corporate social responsibility and environmental protection awareness have also attracted more and more attention from the management, resulting in the phenomenon of “greening”. However, the research shows that the actual performance of environmental responsibility can promote the steady development of enterprises and bring long-term benefits to enterprises. From corporate environmental responsibility to its social responsibility and corporate governance theory, it covers three aspects of ESG research, and is also the different direction of ESG research abroad. On the whole, both domestic and foreign studies tend to pay more attention to issues related to corporate financial performance, corporate value, and the environment, social responsibility and corporate governance, showing a thinking process from multiple perspectives and different aspects, including theoretical discussion and quantitative analysis. Future related research will also be more in-depth, which will have different degrees of impact on the development of enterprises and the performance of social responsibility.

2. Case enterprise introduction

The predecessor of Shanxi J Co., Ltd. was the Preparatory Office of Shanxi fertilizer plant, which was founded in 1969. The first old coke oven was built in 1979. In 1996, it was reorganized into Shanxi Coking Group Co., Ltd. and listed on the Shanghai Stock Exchange on August 8 of the same year. In 2004, it restructured with Shanxi Coking Coal Group and became one of the wholly-owned subsidiaries of Shanxi coking coal. Shanxi coking is a comprehensive coal utilization enterprise that retorts coal, produces coke, and recycles and deep processes coking by-products. It is one of the 82 pilot enterprises of circular economy in China, a “green factory” recognized by the Ministry of industry and information technology of the people’s Republic of China, and an advantageous enterprise for key development in Shanxi Province. It is a listed company in the coking industry established by raising funds with the approval of the people’s Government of Shanxi Province. It was listed on the Shanghai Stock Exchange on August 8th, 1996^[8].

At present, Shanxi J company has 6 JN60 coke ovens with a design capacity of 3.546 million tons/year, supporting 300000 tons/year coal tar processing, 350000 tons/year methanol, 100000 tons/year benzene refining, 80000 tons/year carbon black, and mainly produces 45 kinds of products such as coke, ammonium sulfate, industrial naphthalene, asphalt, anthracene oil, wash oil, phenols, carbon black, methanol, benzene^[9]. Among them, the leading product metallurgical coke was rated as “China’s top quality product” and “Shanxi’s landmark brand product”, and coking benzene, ammonium sulfate, industrial naphthalene and industrial methanol were rated as “Shanxi’s brand products”. The company has 8 production plants and 5 auxiliary production units including coking plant, coal reserve plant, chemical product recovery plant, public auxiliary plant, tar processing plant, benzene refining plant, methanol plant and carbon black plant^[10]. The company holds 75.03% of the shares of Shanxi Coking Coal Feihong Chemical Co., Ltd. and 40% of the shares of Shanxi delixin Electronic Technology Co., Ltd; 49% shares of China Coal Huajin Group Co., Ltd.^[10].

Its core subsidiary, Shanxi J Co., Ltd., was established in 1996 and has been deeply engaged in the field of coal chemical industry. At the initial stage, due to the market demand of the state and Shanxi for coal and related products, certain business achievements were made. However, with the development of the times and the increasingly obvious disadvantages of coal pollution, the state’s increasingly strict management and mining of Shanxi’s coal, coupled with the price changes in the international energy market and the influence of various complex factors, in recent years, the business situation of Shanxi J company has become increasingly severe. On July 14, when Shanxi J company released the 2025 semi annual performance forecast, the market was not surprised to see a deficit. The coking giant with a history of nearly 60 years estimated that the net profit loss attributable to the owner of the parent company in the first half of the year was 72.7692 million yuan to 87.2692 million yuan, in sharp contrast to the profit of 184million yuan in the same period of the previous year. The company has an annual capacity of 3.6 million tons of coke, with more than 5500 employees, and the actual controller is the state owned assets supervision and Administration Commission of Shanxi Province. Despite the halo of “key protected and advantageous enterprises in Shanxi Province”, it has been hard to escape the impact of the downward cycle of the industry in recent years. On the one hand, due to the continuous decline of upstream and downstream prices, the coke market demand is insufficient,

which is unfavorable to the development of enterprises; On the other hand, investment income, operating performance and other indicators continued to decline, unable to make up for cost losses. The realistic pressure and situation faced by Shanxi J company makes the enterprise have to seek transformation and development. Under the background of “double carbon” policy, as a coking enterprise, the pursuit of green operation and low-carbon transformation has become increasingly required by the enterprise.

From the main business income and proportion in 2024 listed in the table below ^[11], Shanxi J company mainly operates in 20 main product categories, including coke, coking toluene, pure benzene and heavy benzene, as well as raw materials, production products and additional products related to the coking industry, such as methanol, carbon black, pure benzene and heavy benzene. The main business income of coking products accounted for 99.44%, and coke as the main product accounted for more than 50%.

Table 2-1 Main Business Income and Proportion of Shanxi J Company in 2024

Main business category	Amount (in billions of yuan)	proportion
coking	74.65	99.44%
coke	51.54	76.65%
methanol	4.55	6.06%
Coking toluene	0.72	0.95%
carbon black	3.79	5.05%
Coking xylene	0.24	0.32%
Pure benzene	4.74	6.32%

Source: Annual Report of Shanxi J Company in 2024

Table 2-2 Main Product Categories of Shanxi Coking

category	Main products
Coking industry products	Coke, methanol, carbon black, pure benzene, modified solid pitch, industrial naphthalene liquid, modified liquid pitch, coking toluene, methylnaphthalene oil, heavy benzene, ammonium sulfate, dephenolized phenol oil, coking xylene, non-aromatic hydrocarbons, wash oil, light oil, neutral phenolic salt, sulfur, molten sulfur, anthracene oil

Source: Shanxi Coking Annual Report 2024

According to a report by Shanxi Evening News on August 31, in the first half of 2025, 41 A-share listed companies in Shanxi achieved a total operating revenue of 257.09 billion yuan, a decrease of 22.86 billion yuan or 8.17% compared to the same period last year ^[11]. The coal industry suffered a performance decline due to the drop in coal prices, and Shanxi Coking was also impacted, with operating revenue and profits declining.

Through analyzing the operating conditions of Shanxi J Company, it can be seen that Shanxi Coking is currently at a critical juncture, which is related to the overall downward trend in the performance of Shanxi’s coal industry. Even so, Shanxi Coking is actively making changes, attempting to help the company out of the predicament by transforming its development mode and exploring sustainable development paths, and the ESG concept will play a very important “role” in this process.

3. Evaluation of ESG Practice Effectiveness of Shanxi J Company

3.1 Entropy-catastrophe progression method for measuring ESG performance

3.1.1 Indicator selection and data preprocessing

3.1.1.1 Selection of indicators

When using the catastrophe progression method, we must select appropriate metrics to evaluate ESG performance. These metrics can be divided into three layers: the first layer includes environment, society, and corporate governance (abbreviated as E, S, G); the second layer is further refined into four subcategories, namely financial performance, business outcomes, development capability, and management innovation; the third layer is at a more specific micro-level, such as specific

corporate behaviors. The indicators selected in this study are mainly based on two points: firstly, the importance of ESG issues in Shanxi J Company; secondly, the significance of Shanxi J Company's performance in these issues and its specific action plans.

In this study, it was found that the relevant information disclosed by Shanxi J Company on its issues is quite limited, thus necessitating the selection of other sustainable information disclosures. The table below presents the significant ESG issues disclosed by Shanxi J Company.

Table 3-1 Important Topics of Shanxi J Company

	Environmental Issues	Social responsibility issues	Corporate governance issues
2023-2024	Addressing climate change, environmental compliance management, energy utilization, pollutant emissions, waste disposal, ecosystem and biodiversity conservation, water resource utilization, and circular economy	Safe production and operation, safeguarding employee rights and interests, promoting employee growth, ensuring the safety and quality of products and services, rural revitalization, social contribution, supply chain security, treating small and medium-sized enterprises equally, data security, and customer privacy protection	Compliance management, corporate governance, anti-commercial bribery and anti-corruption, due diligence, anti-unfair competition, and technology ethics

(1) Environmental indicators

From the ESG report of Shanxi J Company, it can be seen that the enterprise places greater emphasis on environmental compliance management and addressing climate change. Based on this, this article categorizes environmental indicators into two major categories: addressing climate change and emissions, and environmental compliance management and energy utilization. The specific indicators are listed in the table below:

Table 3-2 Environmental Indicators

First-level indicator	Secondary indicator	Third-level indicator	Indicator type
Environment	Addressing climate change and emissions	Total greenhouse gas emissions	Negative indicator
		total exhaust emissions	Negative indicator
	Environmental compliance management and energy utilization	Total volume of industrial wastewater	Negative indicator
		Amount of environmental administrative penalties during the reporting period	Negative indicator
		Comprehensive energy consumption	Negative indicator

(2) Social responsibility indicators

Based on the issues of importance and ESG strategy, the main focus of Shanxi J Company's ESG practice in 2024 is product and service safety and quality. Referring to the actual practices of Shanxi Coking in social responsibility, the indicators are selected as follows:

Table 3-3 Social Responsibility Indicators

First-level indicator	Secondary indicator	Third-level indicator	Indicator type
Social responsibility	Product and service safety and quality	Number of product violation incidents	Negative indicator
		Employee training pass rate	Positive indicator
	Employee rights and growth	Annual training expenditure amount	Positive indicator
		Labor contract signing rate	Positive indicator
	Supply Chain and Innovation	Full-process electronic bidding and procurement coverage rate	Positive indicator
		R&D expenses	Positive indicator

(3) Corporate governance indicators

In the corporate governance issues disclosed by Shanxi Coking, the main focus is on compliance management and internal

integrity supervision. The classification of this indicator is also based on this, as shown in the table below:

Table 3-4 Corporate Governance Indicators

First-level indicator	Secondary indicator	Third-level indicator	Indicator type
Corporate governance	Compliance Management	Identify contract management risks	Negative indicator
		Formulate risk prevention and control measures	Positive indicator
	Integrity and Investor Relations	Number of regulatory training sessions	Positive indicator
		Customer and investor response rate	Positive indicator

3.1.1.2 Data sources

Based on the data disclosed in the Environmental, Social, and Corporate Governance (ESG) report of Shanxi J Company from 2023 to 2024, the ESG evaluation index data of Shanxi J Company is summarized as follows:

Table 3-5 Environmental Indicators

Secondary indicator	Third-level indicator	2023	2024
Addressing climate change and emissions E_1	Total greenhouse gas emissions E_{11}	206.89	197.59
	Total exhaust emissions E_{12}	332.85	334.47
	Total volume of industrial wastewater E_{13}	3987974	1234800
Environmental compliance management and energy utilization E_2	Amount of environmental administrative penalties during the reporting period E_{21}	0	0
	Comprehensive energy consumption E_{22}	120.8	112.68

Table 3-6 Social Responsibility Indicators

Secondary indicator	Third-level indicator	2023	2024
Product and service safety and quality S_1	Number of product violation incidents S_{11}	0	0
	Pass rate of employee safety education and training S_{21}	100%	100%
Employee rights and growth S_2	Annual training expenditure amount S_{22}	603.79	646
	Labor contract signing rate S_{23}	100%	100%
Supply chain and innovation S_3	Full-process electronic bidding and procurement coverage rate S_{31}	100%	100%
	R&D expenses (10,000 yuan) S_{32}	9528	9670.9

Table 3-7 Corporate Governance Indicators

Secondary indicator	Third-level indicator	2023	2024
Compliance Management G_1	Identify contract management risks G_{11}	30	34
	Formulate risk prevention and control measures G_{12}	60	82
Integrity and Investor Relations G_2	Number of regulatory training sessions G_{21}	16	4
	Customer and investor response rate G_{22}	100%	100%

3.1.1.3 Data preprocessing

Due to the different units and characteristics of ESG indicators, which encompass both qualitative and quantitative, positive and negative indicators, normalization is necessary to facilitate calculation and avoid bias^[12].

If it is a positive indicator, the normalization formula is:

$$X_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (4-1)$$

If it is a negative indicator, the normalization formula is:

$$X_i = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}} \quad (4-2)$$

The resulting dimensionless results are presented in the table below:

Table 3-8 Non-dimensionalized Data

After determining the weights, it is necessary to select the corresponding mutation model based on the number of indicators. Given the unique characteristics of ESG evaluation indicators, the mutation level values in this paper are calculated using the average value^[12].

The mutation model corresponding to the ESG evaluation index system of Shanxi J Company is shown in the table below:

Table 3-8 Non-dimensionalized Data

Third-level indicator	2023	2024
Total greenhouse gas emissions E ₁₁	0.00	1.00
Total exhaust emissions E ₁₂	1.00	0.00
Total volume of industrial wastewater E ₁₃	0.00	1.00
Amount of environmental administrative penalties during the reporting period E ₂₁	0.00	1.00
Comprehensive energy consumption E ₂₂	0.00	1.00
Number of product violation incidents S ₁₁	0.00	0.00
Pass rate of employee safety education and training S ₂₁	0.00	0.00
Annual training expenditure amount S ₂₂	0.00	14.30
Labor contract signing rate S ₂₃	0.00	0.00
Full-process electronic bidding and procurement coverage rate S ₃₁	0.00	0.00
R&D expenses (10,000 yuan) S ₃₂	0.00	66.68
Identify contract management risks G ₁₁	1.00	0.00
Formulate risk prevention and control measures G ₁₂	0.00	2.73
Number of regulatory training sessions G ₂₁	1.00	1.25
Customer and investor response rate G ₂₂	0.00	0.00

3.2 Determining weights using the entropy method

The principle of determining weights using the entropy method is based on the dispersion degree of indicators. Assuming there are m samples and n indicators in the sample, then X_{ij} represents the j-th indicator of the i-th sample (i=1,2,3, ...,m; j=1,2,3 ...,n)^[12].

Based on this, the data matrix of indicators is derived, and the weights are finally calculated through four steps: normalization of the dataset, calculation of the proportion of the i-th sample under the j-th indicator, calculation of the entropy value of the j-th indicator, and calculation of the information difference coefficient of the j-th indicator^[12]:

$$w_j = \frac{d_j}{\sum_{j=1}^m d_j} \quad (4-3)$$

The weights of the ESG evaluation indicators calculated in this article are shown in the table below:

Table 3-9 Weight calculation results using the entropy method

Evaluation index	Information entropy value e	Information utility value d	Weight (%)
Pass rate of employee safety education and training S ₂₁	0	1	6.667
Annual training expenditure amount S ₂₂	0	1	6.667
Labor contract signing rate S ₂₃	0	1	6.667
Full-process electronic bidding and procurement coverage rate S ₃₁	0	1	6.667
R&D expenses (10,000 yuan) S ₃₂	0	1	6.667
Formulate risk prevention and control measures G ₁₂	0	1	6.667
Number of regulatory training sessions G ₂₁	0	1	6.667
Customer and investor response rate G ₂₂	0	1	6.667
Total greenhouse gas emissions E ₁₁	0	1	6.667
Total exhaust emissions E ₁₂	0	1	6.667
Total volume of industrial wastewater E ₁₃	0	1	6.667
Amount of environmental administrative penalties during the reporting period E ₂₁	0	1	6.667
Comprehensive energy consumption E ₂₂	0	1	6.667
Number of product violation incidents S ₁₁	0	1	6.667
Identify contract management risks G ₁₁	0	1	6.667

Based on the results in the table above, the weights of secondary and primary indicators can be derived from tertiary indicators through stepwise aggregation. The final weight results of the indicator system are presented in the table below:

Table 3-10 Weightings of Indicators in the ESG Evaluation System

Primary indicator weight (%)	Secondary indicator weight (%)	Tertiary indicator weight (%)
Environment	Addressing climate change and emissions E ₁	Total greenhouse gas emissions E ₁₁ 6.667
		Total exhaust emissions E ₁₂ 6.667
		Total volume of industrial wastewater E ₁₃ 6.667
	Environmental compliance management and energy utilization E ₂	Amount of environmental administrative penalties during the reporting period E ₂₁ 6.667
		Comprehensive energy consumption E ₂₂ 6.667
33.335	20.001	
	13.334	

Primary indicator weight (%)		Secondary indicator weight (%)		Tertiary indicator weight (%)	
Social responsi- bility	40.002	Product and service safety and quality S ₁	6.667	Number of product violation incidents S ₁₁	6.667
				Pass rate of employee safety education and training S ₂₁	6.667
		Employee rights and growth S ₂	20.001	Annual training expenditure amount S ₂₂	6.667
				Labor contract signing rate S ₂₃	6.667
				Full-process electronic bidding and pro- curement coverage rate S ₃₁	6.667
		Supply chain and innovation S ₃	13.334	R&D expenses (10,000 yuan) S ₃₂	6.667
				Identify contract management risks G ₁₁	6.667
				Formulate risk prevention and control measures G ₁₂	6.667
Corporate gover- nance	26.668	Compliance Management G ₁	13.334		
				Number of regulatory training sessions G ₂₁	6.667
		Integrity and Investor Relations G ₂	13.334	Customer and investor response rate G ₂₂	6.667

3.3 Catastrophe-level numerical calculation process

After determining the weights, it is necessary to select the corresponding mutation model based on the number of indicators. According to the unique characteristics of ESG evaluation indicators, the mutation level values in this paper are calculated using the average value^[12].

The mutation model corresponding to the ESG evaluation index system of Shanxi J Company is shown in the table below:

Table 3-11 Corresponding mutation models for evaluation indicators

Evaluation index	Number of control variables	Catastrophe model
Addressing climate change and emissions E_1	3	Swallow-tailed
Environmental compliance management and energy utilization E_2	2	Cusp
Product and service safety and quality S_1	1	Foldable
Employee rights and growth S_2	3	Swallow-tailed
Supply chain and innovation S_3	2	Cusp
Compliance Management G_1	2	Cusp
Integrity and Investor Relations G_2	2	Cusp
Environment	2	Cusp
Social responsibility	3	Swallow-tailed
Corporate governance	2	Cusp
Overall ESG performance	3	Swallow-tailed

Based on the induction of the aforementioned mutation model, the mutation level values are calculated according to the normalization formula and indicator weights, with separate calculations for primary and secondary indicators.

3.4 ESG Practice Evaluation Results of Shanxi J Company

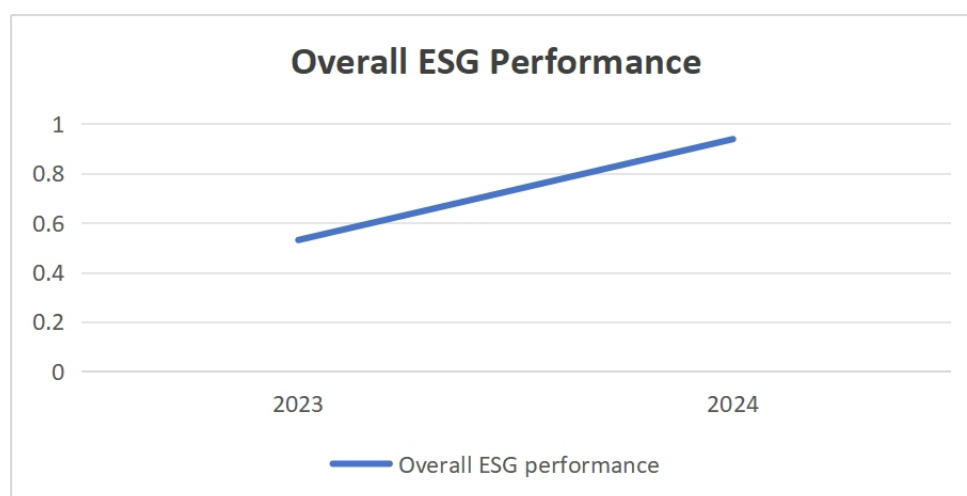
Based on the results mentioned in the previous text, the mutation level values of each primary and secondary indicator for 2023 and 2024 were calculated in sequence, and finally, the overall ESG performance was calculated. The calculation results are shown in the table below:

Table 3-12 Numerical results of mutation levels for ESG secondary indicators

Indicator	2023	2024
Addressing climate change and emissions E_1	0.3333	0.6667
Environmental compliance management and energy utilization E_2	0.0000	1.0000
Product and service safety and quality S_1	0.0000	0.0000
Employee rights and growth S_2	0.0000	0.8090
Supply chain and innovation S_3	0.0000	2.2075
Compliance Management G_1	0.5000	0.6988
Integrity and Investor Relations G_2	0.5000	1.2801

Table 3-13 Results of mutation level values for ESG primary indicators

Indicator	2023	2024
Environment	0.2886	0.9082
Social responsibility	0.0000	0.7338
Corporate governance	0.7504	0.9608
Overall ESG performance	0.5305	0.9383



From the overall ESG score of Shanxi J company in 2023 and 2024, it shows a trend of listing. From 0.5 to 0.9, the most obvious increase is in social responsibility. From 0 to 0.7, the increase in environment and corporate governance is slightly lower. On the whole, in the two years of ESG practice in Shanxi J company, great progress and achievements have been made in all aspects, and the overall performance of ESG shows an upward trend.

In 2024, Shanxi J company released the environmental, social and corporate governance report for the first time. Due to the lack of relevant data and insufficient indicators, the score of social responsibility was low. In the disclosure in 2025, relevant information has been improved and the disclosure at the environmental level has been listed separately. Although the sustainable information disclosed by Shanxi J company is relatively limited and the time dimension is relatively short, the enterprise has shown its strong strength in ESG, and all aspects have improved in a short year, and the enterprise has paid

more and more attention to ESG work.

4. research conclusions and suggestions

4.1 research conclusion

According to the evaluation data of entropy catastrophe progression method, the ESG performance of Shanxi J company from 2023 to 2024 showed a certain level. In particular, it achieved a significant increase in 2024, and the overall score showed an upward trend, especially in the field of corporate governance, which showed that enterprises attached importance to corporate governance. In addition, Shanxi J company has also taken corresponding governance measures at all levels, and environmental and social responsibility are also being improved and put forward. In the evaluation system, social responsibility is regarded as the core concern of Shanxi J company, accounting for the largest proportion. Especially in terms of product services and supply chain management, they have played a vital role in the operation of Shanxi J company, and have become the ESG topics of most concern to stakeholders.

4.2 suggestions for enterprise development

The achievements of Shanxi J company in ESG practice also need to be strengthened and improved in terms of ESG information transparency, sustainable information disclosure completeness and strengthening the digital level of enterprises. For this, the following development suggestions are made:

(1) Improve transparency and completeness of sustainable disclosure

Although the sustainable information disclosure time of Shanxi J company is relatively short, it still shows its potential in sustainable development. Especially since 2024, it shows great potential in the performance of enterprise ESG and the development of digital economy. Current research shows that there is an important link between enterprise ESG performance, digital economy and new quality productivity, which is also an interdisciplinary and multi-dimensional display of ESG. In this context, Shanxi J company should continue to improve its initiatives in related fields, especially in the field of digitalization, take the initiative to strengthen completeness, better meet the needs of investors and stakeholders, and show its corporate image.

(2) Strengthen the construction of enterprise culture

Corporate culture is the unique image of corporate values. As the guide of employee behavior norms, it permeates in all aspects, it has a subtle guiding role^[13]. Shanxi J company adheres to the core values of “taking the striver as the foundation and working hard for a long time”, overcomes difficulties, forge ahead, and is committed to high-quality and sustainable development of the enterprise. In the future, Shanxi J company should take ESG as a part of corporate culture, actively practice the concept of sustainable development, and effectively improve the level of ESG disclosure.

(3) Strengthen the coordinated development of environment, society and corporate governance

From the ESG practice results of Shanxi J company, it can be seen that there are differences in the scores of environment, society and corporate governance in the process of sustainable development of the company, and the difference is large. This shows that enterprises do not pay much attention to the collaborative development of dimensions in the process of ESG practice. As a system, ESG should think in an overall and systematic way, strengthen the collaborative development level of enterprise environment, society and corporate governance, and better promote the sustainable development of enterprises.

Funding

No

Conflict of Interests

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

Reference

- [1] Liu, X., & Wang, P. (2025). Analysis of differences in ESG research systems at home and abroad and changes in frontier hotspots. *Journal of Finance and Accounting*, 46(04), 116–123.
- [2] Li, Z., & Li, B. (2025). Compilation and integration of social responsibility reports and ESG reports of state-owned enterprises. *Quality and Certification*, (07), 34–37.

- [3] Li, E., Wang, B., & Mao, Q. (2025). ESG report assurance: Current status, trends, and challenges. *Quality and Certification*, (07), 45–48.
- [4] Guo, X. (2025). Global perspective: China's ESG development 2024. *Future and Development*, 49(07), 37–43.
- [5] Wang, D., & Huang, J. (2021). *ESG theory and practice*. Economic Management Press.
- [6] Bansal, P., & Roth, K. (2000). Why companies go green: A model of ecological responsiveness. *Academy of Management Journal*, 43(4), 717–736.
- [7] Henriques, I., & Sadosky, P. (1999). The relationship between environmental commitment and managerial perceptions of stakeholder importance. *Academy of Management Journal*, 42(1), 87–99.
- [8] Shanxi Coking. (2023). 2023 Environmental, Social, and Corporate Governance (ESG) Report [Report]. Shanxi Coking Group Co., Ltd.
- [9] Shanxi Coking Group Co., Ltd. (2020). [Publisher information]. *Shanxi Chemical Industry*, 40(05), 2.
- [10] Shanxi Coking. (2024). 2024 Environmental, Social, and Corporate Governance (ESG) Report [Report]. Shanxi Coking Group Co., Ltd.
- [11] Bai, J. (2023). Research on the development strategy of Shanxi Coking Group [Master's thesis]. Lanzhou University of Technology.
- [12] Cheng, J. (2024). Research on the relationship and influence path between ESG performance and corporate value [Master's thesis]. Inner Mongolia University of Finance and Economics.
- [13] Wang, C. (2024). Research on the impact of ESG performance on corporate value in coal enterprises [Master's thesis]. Henan Agricultural University.

A Bilevel GA-PSO-MILP Framework for Capacity Planning and Day-Ahead Scheduling of an Integrated Electricity-Heat-Cooling Energy System

Mengzhe Yu¹, Mingshen Xu^{1*}, Bomian Cheng²

1. North China Electric Power University, Baoding, Hebei, 071003, China

2. Chengdu University of Technology, Chengdu, Sichuan, 610059, China

*Corresponding author: Mingshen Xu, 3165865378@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 proposes a bilevel planning–operation framework for an integrated electricity–heat–cooling energy system comprising photovoltaic generation, wind generation, a gas-turbine-based combined heat and power unit, a ground source heat pump, a gas boiler, battery storage, thermal storage, an absorption chiller, and an electric chiller. The upper level determines the installed capacities of major devices, while the lower level performs day-ahead coordinated dispatch over four representative days. To preserve the engineering logic of the original optimization program while improving analytical transparency and reproducibility, the planning layer is formulated as a GA-PSO-based search procedure and the coupled planning–operation problem is further expressed as an equivalent mixed-integer linear benchmark model. A complete mathematical formulation is provided, including explicit decision-variable definitions, energy-conversion relationships, storage dynamics, logical constraints, and a comprehensive notation table. Using the representative-day data embedded in the program package, the optimal configuration is obtained as 800 kWh battery storage, 900 kW photovoltaic, 600 kW wind, 1200 kW gas turbine, 300 kW ground source heat pump, 200 kW gas boiler, 600 kWh thermal storage, 479.08 kW absorption chiller, and 150 kW electric chiller. The corresponding daily equivalent annualized investment cost is 6286.79, the day-ahead composite operating objective is 77400.43, and the ex-post composite operating objective under realized loads is 78763.31. The results indicate that strong renewable deployment, flexible storage scheduling, and coordinated electricity–heat–cooling conversion can substantially improve the economic and operational performance of integrated energy systems.

Keywords: Integrated Energy System; Bilevel Optimization; Capacity Planning; Day-Ahead Scheduling; Ga-Pso; Mixed-Integer Linear Programming

Published: Apr 5, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1229>

1. Introduction

Integrated energy systems provide a physically coupled platform for the coordinated production, conversion, storage, and utilization of multiple energy carriers, including electricity, heat, cooling, and gas. By exploiting complementary interactions among different conversion devices and demand types, such systems can improve renewable-energy utilization, reduce operating cost, and enhance local energy flexibility. Energy hub theory established the conceptual basis for multi-carrier coordination ^[1-2], and subsequent studies on distributed multi-generation, building energy systems, and multi-energy micro-

grids gradually extended this concept toward practical planning and operation problems ^[3-5]. At broader spatial scales, integrated electricity–heat–gas and electricity–heat–cooling systems have been studied from infrastructure, planning, dispatch, and policy perspectives ^[6-9].

Recent research has increasingly focused on more detailed planning and operational models for integrated energy systems. Optimization frameworks have been proposed to represent technology coupling, demand response, infrastructure interaction, and uncertainty across different carriers ^[10-12]. Two-stage, robust, and decomposition-based methods have improved the treatment of uncertain loads and renewable availability ^[13-14]. In parallel, resilience-oriented planning, carbon-aware expansion, and integrated electricity–heat design methods have further enriched the methodological landscape ^[15-18]. Despite these advances, two practical issues remain common in engineering-oriented studies. First, the planning layer and the operational layer are often described only at a conceptual level, which makes it difficult to identify the exact coupling between investment decisions and day-ahead dispatch. Second, program-based implementations frequently contain complete computational logic but insufficient mathematical exposition, especially regarding notation, logical constraints, and equivalent linear reformulations.

For upper-level capacity planning, heuristic search remains attractive because the capacity mix of integrated energy systems is affected by nonconvex equipment interactions, long-term investment trade-offs, and feedback from lower-level operation. The methodological roots of this approach can be traced to genetic algorithms and evolutionary search ^[19], as well as particle swarm optimization ^[20]. For lower-level coordinated dispatch, algebraic modeling platforms and commercial mixed-integer optimization solvers are widely used because they can efficiently handle time-coupled scheduling constraints and logical operating states ^[21-22]. The optimization code supporting this study follows exactly this combination: a GA–PSO planning layer is coupled with a YALMIP–GUROBI operational dispatch model over four representative days.

Motivated by the need for a more rigorous and publication-ready presentation, this paper reorganizes the original engineering model into a complete bilevel planning–operation manuscript. The contribution of the paper is fourfold. First, the coupled electricity, heat, cooling, gas, and storage interactions are systematically reformulated as a bilevel capacity-planning and day-ahead scheduling model. Second, the upper-level GA–PSO search philosophy is retained, while the full planning–operation linkage is further rewritten as an equivalent mixed-integer linear benchmark model to improve reproducibility and analytical clarity. Third, a unified notation system and explicit parameter specification are provided to eliminate ambiguity in variables and constraints. Fourth, the numerical case study interprets the optimal capacity mix and the representative-day dispatch behavior from the perspective of renewable integration, CHP-led coupling, storage flexibility, and dual cooling technologies. The remainder of this paper is organized as follows. Section 2 introduces the integrated electricity–heat–cooling system and the bilevel decision structure. Section 3 presents the complete mathematical formulation. Section 4 describes the solution methodology, including the GA–PSO procedure and the exact MILP benchmark reformulation. Section 5 reports the case-study settings, and Section 6 discusses the numerical results. Section 7 concludes the paper.

2. System Description and Problem Statement

2.1 Integrated Electricity–Heat–Cooling System Architecture

The system considered in this paper contains three tightly coupled demand categories: electricity demand, heat demand, and cooling demand. The electricity subsystem consists of photovoltaic generation, wind generation, the electric output of a gas-turbine-based combined heat and power (CHP) unit, battery storage, the electrical input of the ground source heat pump, the electrical input of the electric chiller, and bidirectional power exchange with the external grid. The heat subsystem consists of CHP thermal recovery, the heat supplied by the ground source heat pump, a gas boiler, thermal storage, and the heat consumed by the absorption chiller. The cooling subsystem is jointly served by the electric chiller and the absorption chiller. Natural gas is purchased from the external gas network and is consumed by the CHP unit and the gas boiler.

The operational scheduling horizon is 24 hours for each representative day. Long-term planning is reflected through daily equivalent annualized investment costs. Let d denote the representative-day index and t the hourly index. The upper level determines the installed capacity vector

$$\mathbf{X} = [\mathbf{X}_{es}, \mathbf{X}_{pv}, \mathbf{X}_{wt}, \mathbf{X}_{gt}, \mathbf{X}_{hp}, \mathbf{X}_{gb}, \mathbf{X}_{hs}, \mathbf{X}_{ac}, \mathbf{X}_{ec}]^T. \quad (1)$$

where each element represents the installed size of a major technology. Given X , the lower level optimizes the day-ahead dispatch of all energy carriers and controllable devices.

2.2 Bilevel Planning–Operation Structure

The interaction between planning and operation is depicted by a master–subproblem structure. The upper level selects device capacities by balancing capital expenditure against the operating performance induced by those capacities. The lower level receives a fixed capacity vector and schedules the system over each representative day so as to minimize a composite operating objective that includes direct cost and environmental penalty. This structure can be written compactly as

$$\min_{X \in \mathcal{X}} F(X) = C^{\text{inv}}(X) + \sum_{d=1}^4 \omega_d \Phi_d(X) \quad (2)$$

where X denotes the feasible planning set, ω_d is the weight of representative day d , c^{inv} is the daily equivalent annualized investment cost, and $\Phi_d(X)$ is the optimal lower-level objective value corresponding to representative day d . In the present case study, all representative-day weights are set to one, consistent with the aggregation logic adopted in the optimization code.

For a fixed capacity vector X , the lower-level problem on representative day d is

$$\Phi_d(X) = \min_{y_d \in \mathcal{Y}_d(X)} (C_d^{\text{dir}} + C_d^{\text{em}}) \quad (3)$$

where y_d collects all operational decision variables and $\mathcal{Y}_d(X)$ denotes the day-ahead feasible region induced by energy-balance equations, conversion relationships, storage dynamics, logical operating states, and device-capacity limits.

3. Mathematical Formulation

3.1 Upper-Level Planning Model

The admissible range of each planning variable is

$$X_j \leq x_j \leq \bar{x}_j, \quad \forall j \in \{\text{es, pv, wt, gt, hp, gb, hs, ac, ec}\} \quad (4)$$

where x_j and \bar{x}_j are the lower and upper capacity bounds, respectively. The daily equivalent annualized investment cost is calculated through the capital recovery factor:

$$\text{CRF}_j = \frac{r(1+r)^{L_j}}{(1+r)^{L_j}-1} \quad (5)$$

$$C^{\text{inv}}(X) = \frac{1}{365} \sum_j \text{CRF}_j c_j^{\text{cap}} x_j \quad (6)$$

where r is the discount rate, L_j is the service life of technology j , and c_j^{cap} is the unit capital cost.

Equation (6) maps long-term investment into an equivalent daily cost so that it can be aggregated with representative-day operating terms in the upper-level objective (2).

3.2 Lower-Level Direct Operating Objective

For representative day d , the direct operating expenditure consists of electricity trading cost, gas purchase cost, and variable operation and maintenance cost:

$$\begin{aligned} C_d^{\text{dir}} = & \sum_{t=1}^{24} \lambda_{d,t}^e (p_{d,t}^{\text{buy}} - p_{d,t}^{\text{sell}}) + \sum_{t=1}^{24} \lambda_{d,t}^g g_{d,t}^{\text{buy}} \\ & + \sum_{t=1}^{24} [c_{\text{pv}}^{\text{om}} p_{d,t}^{\text{pv}} + c_{\text{wt}}^{\text{om}} p_{d,t}^{\text{wt}} + c_{\text{gt}}^{\text{om}} p_{d,t}^{\text{gt}} + c_{\text{es}}^{\text{om}} (p_{d,t}^{\text{ch}} + p_{d,t}^{\text{dis}})] \\ & + c_{\text{hp}}^{\text{om}} p_{d,t} + c_{\text{gb}}^{\text{om}} h_{d,t}^{\text{gb}} + c_{\text{hs}}^{\text{om}} (h_{d,t}^{\text{ch}} + h_{d,t}^{\text{dis}}) + c_{\text{ac}}^{\text{om}} c_{d,t}^{\text{ac}} + c_{\text{ec}}^{\text{om}} p_{d,t}^{\text{ec}} \end{aligned} \quad (7)$$

In addition to direct expenditure, the model includes an environmental penalty term:

$$C_d^{\text{em}} = \sum_{t=1}^{24} [\pi_{\text{grid}} (p_{d,t} - p_{d,t}^{\text{sell}}) + \pi_{\text{gt,e}} p_{d,t}^{\text{gt}} + \pi_{\text{gt,h}} h_{d,t}^{\text{gt}} + \pi_{\text{gb}} h_{d,t}^{\text{gb}}] \quad (8)$$

The coefficients π_{grid} , $\pi_{\text{gt,e}}$, $\pi_{\text{gt,h}}$, and π_{gb} are the penalty factors used in the original program. They allow the dispatch model to internalize the environmental implications of grid electricity, CHP operation, and boiler heat production.

3.3 Energy-Conversion Relationships

The gas-turbine-based CHP unit converts natural gas into both electricity and heat. Let $g_{d,t}^{\text{gt}}$ denote the gas flow consumed by the turbine. The power–heat output relationships are

$$p_{d,t}^{\text{gt}} = 0.3 H_g g_{d,t}^{\text{gt}}, \quad h_{d,t}^{\text{gt}} = 0.4 H_g g_{d,t}^{\text{gt}} \quad (9)$$

where $H_g = 9.78$ kWh/Nm³ is the lower heating value of natural gas. Equation (9) implies that the electricity-to-heat output ratio of the CHP unit is fixed, so the dispatch of the turbine simultaneously affects both the electric and thermal subsystems.

The ground source heat pump, electric chiller, and absorption chiller are modeled through fixed conversion coefficients:

$$h_{d,t}^{hp} = \text{COP}_{hp} p_{d,t}^{hp}, c_{d,t}^{ec} = \text{COP}_{ec} p_{d,t}^{ec}, c_{d,t}^{ac} = \text{COP}_{ac} h_{d,t}^{ac}, \quad (10)$$

with $\text{COP}_{hp} = 3.5$, $\text{COP}_{ec} = 4.2$, and $\text{COP}_{ac} = 1.2$. The gas boiler is represented as

$$g_{d,t}^{gb} = \frac{h_{d,t}^{gb}}{\eta_{gb} H_g}, \quad \eta_{gb} = 0.9 \quad (11)$$

Accordingly, the total gas purchase from the external gas network is

$$g_{d,t}^{buy} = g_{d,t}^{gt} + g_{d,t}^{gb} \quad (12)$$

3.4 Carrier-Balance Equations

The coordinated dispatch must satisfy the electricity, heat, and cooling balances at every hour. The electricity balance is

$$p_{d,t}^{buy} - p_{d,t}^{sell} + p_{d,t}^{pv} + p_{d,t}^{wt} + p_{d,t}^{gt} + p_{d,t}^{dis} - p_{d,t}^{ch} - p_{d,t}^{hp} - p_{d,t}^{ec} = p_{d,t}^L \quad (13)$$

which states that grid exchange, renewable generation, CHP electric output, and battery discharge must cover electric demand as well as the electricity consumption of storage charging, the ground source heat pump, and the electric chiller.

The heat balance is

$$h_{d,t}^{gt} + h_{d,t}^{hp} + h_{d,t}^{gb} + h_{d,t} - h_{d,t}^{ch} - h_{d,t}^{ac} = \hat{H}_{d,t}^L \quad (14)$$

and the cooling balance is

$$c_{d,t}^{ee} + c_{d,t}^{ac} = \hat{C}_{d,t}^L \quad (15)$$

Equations (13)–(15) are the core coupling constraints of the integrated energy system because any change in device dispatch propagates across multiple carriers through the conversion relationships defined above.

3.5 Storage Dynamics and Exact Linear Logic

To improve mathematical transparency, the battery and thermal-storage subsystems are formulated with explicit energy-state variables rather than normalized state-of-charge variables. Let $E_{d,t}^d$ denote the battery energy state. The inter-temporal dynamics are

$$E_{d,t}^{es} = E_{d,t-1}^{es} + \eta_{es} P_{d,t} - \frac{1}{\eta_{es}} p_{d,t}, \quad t = 1, \dots, 24, \quad (16)$$

where $\eta_{es} = 0.96$. The energy-state and power limits are

$$0.1X_{es} \leq E_{d,t}^{es} \leq 0.9X_{es}, \quad (17)$$

$$0 \leq p_{d,t}^{ch} \leq 0.5X_{es}, \quad 0 \leq p_{d,t}^{dis} \leq 0.5X_{es}. \quad (18)$$

The initial and terminal conditions enforce daily cycling consistency:

$$E_{d,0}^{es} = 0.55X_{es}, \quad E_{d,24}^{es} = 0.55X_{es}. \quad (19)$$

To exclude simultaneous charging and discharging, binary variables $u_{d,t}^{ch}$ and $u_{d,t}^{dis}$ are introduced:

$$p_{d,t}^{ch} \leq 0.5X_{es}, \quad p_{d,t}^{ch} \leq 0.5\bar{X}_{es} u_{d,t}^{ch}, \quad (20)$$

$$p_{d,t}^{dis} \leq 0.5X_{es}, \quad p_{d,t}^{dis} \leq 0.5\bar{X}_{es} u_{d,t}^{dis} \quad (21)$$

$$u_{d,t}^{ch} + u_{d,t}^{dis} \leq 1, \quad u_{d,t}^{ch}, u_{d,t}^{dis} \in \{0,1\}. \quad (22)$$

This reformulation is exact because the upper bound x_{es} is known from the planning domain. When a binary state is zero, the corresponding charging or discharging power must also be zero; when the binary state is one, the admissible power is still restricted by the installed capacity x_{es} .

The thermal-storage model is analogous. Let $E_{d,t}^{hs}$ denote the thermal energy state. Then

$$E_{d,t}^{hs} = E_{d,t-1}^{hs} + \eta_{hs} h_{d,t} - \frac{1}{\eta_{hs}} h_{d,t}, \quad t = 1, \dots, 24, \quad (23)$$

with $\eta_{hs} = 0.95$, and

$$0.1X_{hs} \leq E_{d,t}^{hs} \leq 0.9X_{hs}, \quad (24)$$

$$0 \leq h_{d,t}^{ch} \leq 0.5X_{hs}, \quad 0 \leq h_{d,t}^{dis} \leq 0.5X_{hs}, \quad (25)$$

$$E_{d,0}^{hs} = 0.55X_{hs}, \quad E_{d,24}^{hs} = 0.55X_{hs}. \quad (26)$$

The mutually exclusive charging/discharging logic is represented by

$$h_{d,t}^{ch} \leq 0.5X_{hs}, \quad h_{d,t}^{ch} \leq 0.5\bar{X}_{hs} u_{d,t}^{h,ch}, \quad (27)$$

$$h_{d,t}^{dis} \leq 0.5X_{hs}, \quad h_{d,t}^{dis} \leq 0.5\bar{X}_{hs} u_{d,t}^{h,dis}, \quad (28)$$

$$u_{d,t}^{h,ch} + u_{d,t}^{h,dis} \leq 1, \quad u_{d,t}^{h,ch}, u_{d,t}^{h,dis} \in \{0,1\}. \quad (29)$$

3.6 Technology Operating Limits

The renewable outputs are restricted by the exogenous availability profiles:

$$0 \leq p_{d,t}^{pv} \leq \alpha_{d,t}^{pv} X_{pv}, 0 \leq p_{d,t}^{wt} \leq \alpha_{d,t}^{wt} X_{wt}. \quad (30)$$

Bidirectional grid exchange is modeled through binary operating states:

$$0 \leq p_{d,t}^{buy} \leq \bar{P}^{grid} u_{d,t}^{buy}, 0 \leq p_{d,t}^{sell} \leq \bar{P}^{grid} u_{d,t}^{sell} \quad (31)$$

$$u_{d,t}^{buy} + u_{d,t}^{sell} \leq 1, u_{d,t}^{buy}, u_{d,t}^{sell} \in \{0,1\}, \quad (32)$$

Where $\bar{P}^{grid} = 500$ kW is the maximum hourly grid-exchange limit used in the scheduling model.

The CHP unit must satisfy minimum-output and ramping constraints. Since x_{gt} denotes the electric capacity of the gas turbine, the output bound is

$$p_{gt}^{min} \leq p_{d,t} \leq X_{gt}, p_{gt}^{min} = 20 \text{ kW}, \quad (33)$$

which is equivalently expressed in gas-flow form as

$$\frac{p_{gt}^{min}}{0.3H_g} \leq g_{d,t}^{gt} \leq \frac{X_{gt}}{0.3H_g}. \quad (34)$$

The hourly ramping constraint is

$$-40 \leq p_{d,t}^{gt} - p_{d,t-1}^{gt} \leq 40, t = 2, \dots, 24. \quad (35)$$

The remaining controllable devices are bounded by their installed capacities:

$$0 \leq p_{d,t}^{hp} \leq X_{hp}, 0 \leq p_{d,t}^{ec} \leq X_{ec} \quad (36)$$

$$0 \leq h_{d,t}^{gb} \leq X_{gb}, 0 \leq h_{d,t}^{ac} \leq X_{ac}. \quad (37)$$

In the operational logic used in the program, the heat input of the absorption chiller is additionally limited by the CHP thermal output:

$$h_{d,t}^{ac} \leq h_{d,t}^{gt}. \quad (38)$$

This constraint implies that absorption cooling is activated primarily when CHP heat is available, thereby strengthening the electricity–heat–cooling coupling structure of the system.

3.7 Equivalent Single-Level Milp Benchmark

Although the original implementation evaluates lower-level dispatch within a heuristic upper-level search, the complete planning–operation problem can be written as an equivalent single-level MILP benchmark:

$$\min_{X, \{y_d\}_{d=1}^4} C^{inv}(X) + \sum_{d=1}^4 \omega_d (C_d^{dir} + C_d^{em}) \quad (39)$$

subject to (4)–(38) for all representative days and hours. This reformulation is exact under the present modeling assumptions because: (i) all investment terms are linear in installed capacities; (ii) all conversion relationships are linear; (iii) all time-coupled storage equations are linear; and (iv) the only capacity–logic interactions are represented by exact mixed-integer linear constraints using known upper bounds. Therefore, the MILP benchmark provides a deterministic reference solution against which the GA–PSO planning results can be checked.

3.8 Comprehensive Notation Summary

Table 1 summarizes the symbols used in the manuscript.

Table 1: Comprehensive notation table covering all variables used in the manuscript.

Symbol	Definition	Dimension / unit
Sets and indices		
d	representative-day index	$d = 1, \dots, 4$
t	hourly index within one representative day	$t = 1, \dots, 24$
j	technology index	$j \in \{es, pv, wt, gt, hp, gb, hs, ac, ec\}$

Symbol	Definition	Dimension / unit
w_d	weight of representative day d	dimensionless
Planning variables and bounds		
x_{es}	installed capacity of battery energy storage	kWh
x_{pv}	installed capacity of photovoltaic generation	kW
x_{wt}	installed capacity of wind generation	kW
x_{gt}	installed electric capacity of the CHP gas turbine	kW
x_{hp}	installed electrical capacity of the ground source heat pump	kW
x_{gb}	installed thermal capacity of the gas boiler	kW
x_{hs}	installed thermal-storage capacity	kWh
x_{ac}	installed heat-input capacity of the absorption chiller	kW
x_{ec}	installed electrical capacity of the electric chiller lower and upper planning	kW
$\underline{X}_j, \bar{X}_j$	bounds of technology j	kW or kWh
Exogenous profiles and prices		
$\hat{P}_{d,t}^L$	forecast electricity load at representative day d, hour t	kW
$\hat{H}_{d,t}^L$	forecast heat load at representative day d, hour t	kW
$\hat{C}_{d,t}^L$	forecast cooling load at representative day d, hour t	kW
$\hat{P}_{d,t}^L$	realized electricity load used in ex-post evaluation	kW
$\hat{H}_{d,t}^L$	realized heat load used in ex-post evaluation	kW
$\hat{C}_{d,t}^L$	realized cooling load used in ex-post evaluation	kW
$\partial_{d,t}^{pv}$	available photovoltaic output factor	p.u.
$\partial_{d,t}^{wt}$	available wind output factor	p.u.
$\lambda_{d,t}^e$	time-of-use electricity tariff	CNY/kWh
$\lambda_{d,t}^g$	gas purchase price	CNY/Nm ³
Operational variables		
$P_{d,t}^{pv}$	scheduled photovoltaic output	kW
$P_{d,t}^{wt}$	scheduled wind output	kW
$P_{d,t}^{buy}$	grid electricity purchase	kW
$P_{d,t}^{sell}$	grid electricity sale	kW
$u_{d,t}^{buy}, u_{d,t}^{sell}$	binary grid-trading states	{0, 1}

Symbol	Definition	Dimension / unit
$P_{d,t}^{ch} P_{d,t}^{dis}$	battery charging and discharging power	kW
$u_{d,t}^{ch} u_{d,t}^{dis}$	binary battery charging and discharging states	{0, 1}
$E_{d,t}^{es}$	battery energy state	kWh
$g_{d,t}^{gt}$	gas consumption of the CHP gas turbine	Nm ³ /h
$p_{d,t}^{gt}$	thermal output of the CHP gas turbine	kW
$p_{d,t}^{hp}$	electric input of the ground source heat pump	kW
$h_{d,t}^{hp}$	heat output of the ground source heat pump	kW
$p_{d,t}^{ec}$	electric input of the electric chiller	kW
$c_{d,t}^{ec}$	cooling output of the electric chiller	kW
$h_{d,t}^{gb}$	heat output of the gas boiler	kW
$h_{d,t}^{ch} h_{d,t}^{dis}$	thermal-storage charging and discharging power	kW
$u_{d,t}^{h,ch} u_{d,t}^{h,dis}$	binary thermal-storage charging and discharging states	{0, 1}
$E_{d,t}^{hs}$	thermal-storage energy state	kWh
$h_{d,t}^{ac}$	heat input of the absorption chiller	kW
$c_{d,t}^{ac}$	cooling output of the absorption chiller	kW
$g_{d,t}^{gb}$	gas consumption of the gas boiler	Nm ³ /h
$g_{d,t}^{buy}$	total gas purchase	Nm ³ /h
Model parameters and objective terms		
H_g	lower heating value of natural gas	9.78 kWh/Nm ³
η_{es}	battery charging/discharging efficiency	0.96
η_{hs}	thermal-storage charging/discharging efficiency	0.95
η_{gb}	gas-boiler efficiency	0.9
COP_{hp}	coefficient of performance of the ground source heat pump	3.5
COP_{ec}	coefficient of performance of the electric chiller	4.2
COP_{ac}	coefficient of performance of the absorption chiller	1.2
c_j^{cop}	unit capital cost of technology j	CNY/kW or CNY/kWh
c_j^{om}	variable operation and maintenance cost of technology j	CNY/kWh
L_j	service life of technology j	year
r	discount rate	0.08

Symbol	Definition	Dimension / unit
CRF_j	capital recovery factor of technology j	dimensionless
π^{grid}	environmental penalty factor of grid electricity	penalty unit/kWh
$\pi^{\text{gt,e}}, \pi^{\text{gt,h}}$	environmental penalty factors of CHP electric and thermal output	penalty unit/kWh
π^{gb}	environmental penalty factor of boiler heat production	penalty unit/kWh
c^{inv}	daily equivalent annualized investment cost	CNY/day
c_d^{dir}	direct operating expenditure of representative day d	CNY/day
c_d^{em}	environmental penalty term of representative day d	penalty unit/day
Φ_d	optimal lower-level objective value on representative day d	composite unit/day
F	total planning–operation objective	composite unit

4. Solution Methodology

4.1 GA–PSO-Based Upper-Level Search

The upper-level planning problem is solved by a hybrid genetic algorithm and particle swarm optimization (GA–PSO) procedure. Let \mathbf{x}_i^k and \mathbf{v}_i^k denote the position and velocity of particle i at iteration k , respectively. The standard PSO update is

$$\mathbf{v}_i^{k+1} = \omega \mathbf{v}_i^k + c_1 r_1 (\mathbf{p}_i^k - \mathbf{x}_i^k) + c_2 r_2 (\mathbf{g}^k - \mathbf{x}_i^k) \quad (40)$$

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \mathbf{v}_i^{k+1} \quad (41)$$

where \mathbf{p}_i^k is the personal-best position of particle i , \mathbf{g}^k is the global-best position of the swarm, ω is the inertia weight, c_1 and c_2 are learning coefficients, and $r_1, r_2 \sim U(0, 1)$ are independent random numbers. In the present study, each particle represents a candidate capacity vector \mathbf{X} . For each candidate, the lower-level dispatch problem is solved on all representative days, and the resulting operating objective is combined with the investment term to evaluate the fitness value in (2). To strengthen global exploration, crossover and mutation operators are embedded after the PSO position update. Crossover recombines selected capacity coordinates from high-quality particles, while mutation randomly perturbs part of the population to prevent premature convergence. In the computational setting inherited from the original program, the population size is 20, the maximum number of iterations is 30, the inertia weight is 0.8, the two learning factors are both 0.5, the mutation rate is 0.05, and the crossover rate is 0.10.

4.2 Planning–Dispatch Evaluation Mechanism

The logic of the bilevel solution can be described as follows.

First, an initial population of feasible capacity vectors is generated within the planning bounds in (4). Second, for each candidate capacity vector, the day-ahead dispatch problem is solved over all representative days, subject to the full set of energy-balance, storage, and logical operating constraints. Third, the daily equivalent investment term and the weighted sum of representative-day operating terms are aggregated into a scalar fitness value. Fourth, the population is updated using the PSO velocity–position rules, followed by crossover and mutation. Fifth, infeasible particles are projected back into the admissible capacity domain. These steps are repeated until the stopping criterion is met, after which the best particle is reported as the heuristic planning solution.

This mechanism highlights the essential feature of the bilevel structure: the upper-level capacity decision is not evaluated through a simplified proxy objective, but through the exact day-ahead operation that the capacity mix induces. Therefore, the planning result already internalizes the temporal interaction among renewable generation, CHP operation, storage scheduling, grid exchange, and dual cooling technologies.

4.3 Role of the Exact Milp Benchmark

Although the heuristic GA–PSO framework is valuable for preserving the original engineering solution philosophy, a deterministic benchmark is useful for verifying the final planning outcome and for reporting a reproducible numerical optimum. The MILP benchmark in (39) serves exactly this purpose. Once the binary logical states are retained and all nonlinearities are replaced by exact mixed-integer linear relations, the integrated planning–operation problem becomes solvable by a commercial MILP solver. The benchmark therefore complements the heuristic search in two ways.

First, it provides a transparent and solver-independent mathematical statement of the entire model. Second, it allows the reported capacity configuration and cost values to be interpreted as the solution of a rigorously defined optimization problem rather than as the output of a black-box script. In this sense, the GA–PSO model and the MILP benchmark should not be viewed as competing formulations, but as two complementary representations of the same planning philosophy: the former mirrors the practical implementation process, whereas the latter improves theoretical clarity and reproducibility.

5. Case Study Settings

The numerical study uses four representative days with hourly electricity, heat, and cooling load forecasts, together with realized load trajectories for ex-post evaluation. The data package also provides hourly upper envelopes of photovoltaic and wind availability. The time-of-use electricity tariff ranges from 0.35 to 1.35 CNY/kWh, while the gas price ranges from 3.20 to 3.80 CNY/Nm³. The planning bounds extracted from the code are summarized in Table 2.

Table 2: Planning bounds and optimal capacities.

Technology	Symbol	Lower	Upper	Optimal	Unit
Battery energy storage	x_{es}	400.00	800.00	800.00	kWh
Photovoltaic	x_{pv}	600.00	900.00	900.00	kW
Wind turbine	x_{wt}	400.00	600.00	600.00	kW
Gas turbine / CHP	x_{gt}	800.00	1200.00	1200.00	kW
Ground source heat pump	x_{hp}	300.00	500.00	300.00	kW
Gas boiler	x_{gb}	150.00	200.00	200.00	kW
Thermal storage	x_{hs}	400.00	600.00	600.00	kWh
Absorption chiller	x_{ac}	400.00	600.00	479.08	kW
Electric chiller	x_{cc}	100.00	150.00	150.00	kW

Table 3 lists the main techno-economic parameters used in planning and operation.

Table 3: Main techno-economic parameters.

Technology	Capex	Life	Capex unit	O&M	O&M unit
Battery energy storage	1500.00	10	CNY/kWh	0.0018	CNY/kWh
Photovoltaic	4800.00	20	CNY/kW	0.4	CNY/kWh
Wind turbine	6000.00	15	CNY/kW	0.9	CNY/kWh
Gas turbine / CHP	7900.00	15	CNY/kW	0.007	CNY/kWh
Ground source heat pump	1050.00	20	CNY/kW	0.026	CNY/kWh
Gas boiler	900.00	15	CNY/kW	0.0082	CNY/kWh
Thermal storage	150.00	10	CNY/kWh	0.0015	CNY/kWh
Absorption chiller	1200.00	15	CNY/kW	0.03	CNY/kWh
Electric chiller	800.00	15	CNY/kW	0.08	CNY/kWh

The prediction accuracy of the three demand carriers is summarized in Table 4. The mean absolute percentage error remains around 5%, indicating that the forecasts are sufficiently accurate for day-ahead coordination while still leaving non-negligible deviation for ex-post assessment.

Table 4: Forecast-error statistics of the representative-day demand data.

Load type	MAE	RMSE	MAPE (%)
Electricity	29.88	35.86	4.92
Heat	20.12	27.07	4.64
Cooling	22.24	29.75	5.02

Figure 1 shows the representative-day renewable availability, and Figures 2–4 compare the forecast and realized demand profiles.

Figure 1: Representative-day wind and photovoltaic availability profiles.

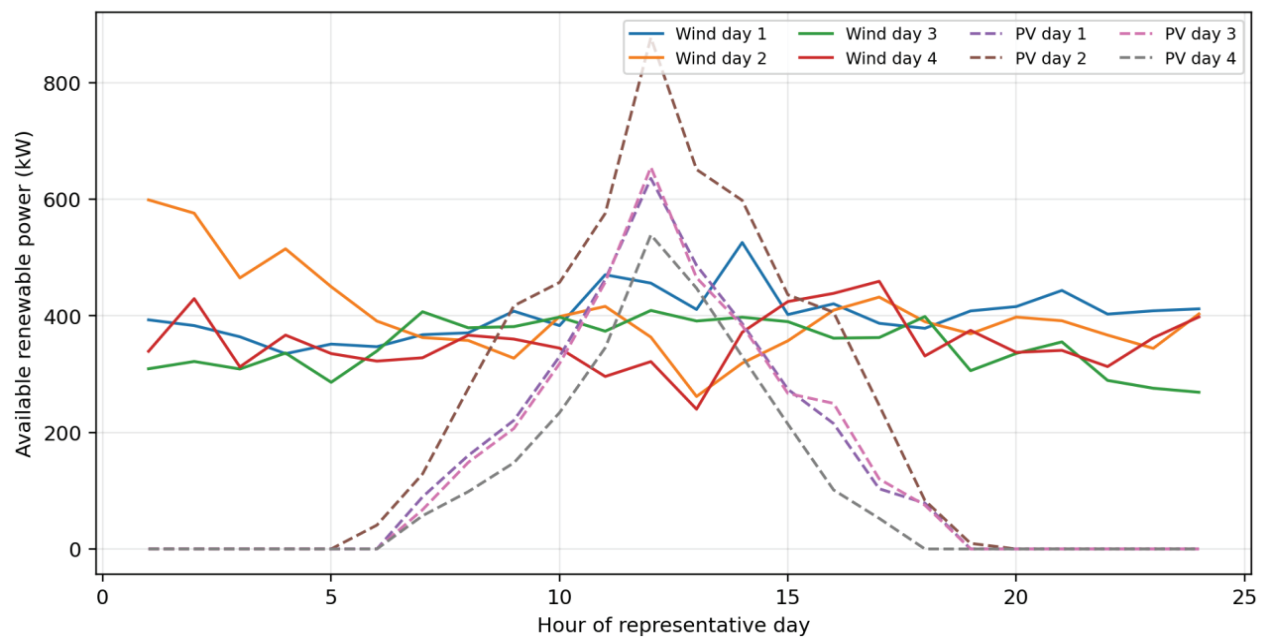


Figure 2: Forecast and realized electricity loads.

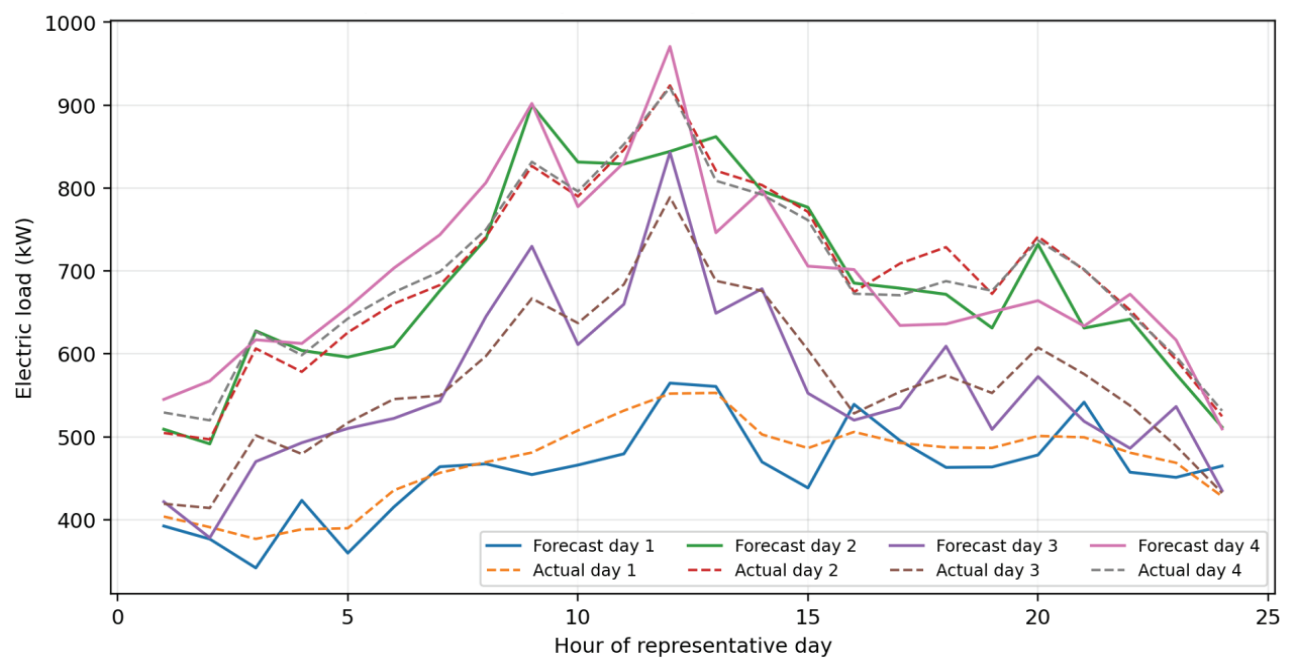


Figure 3: Forecast and realized heat loads.

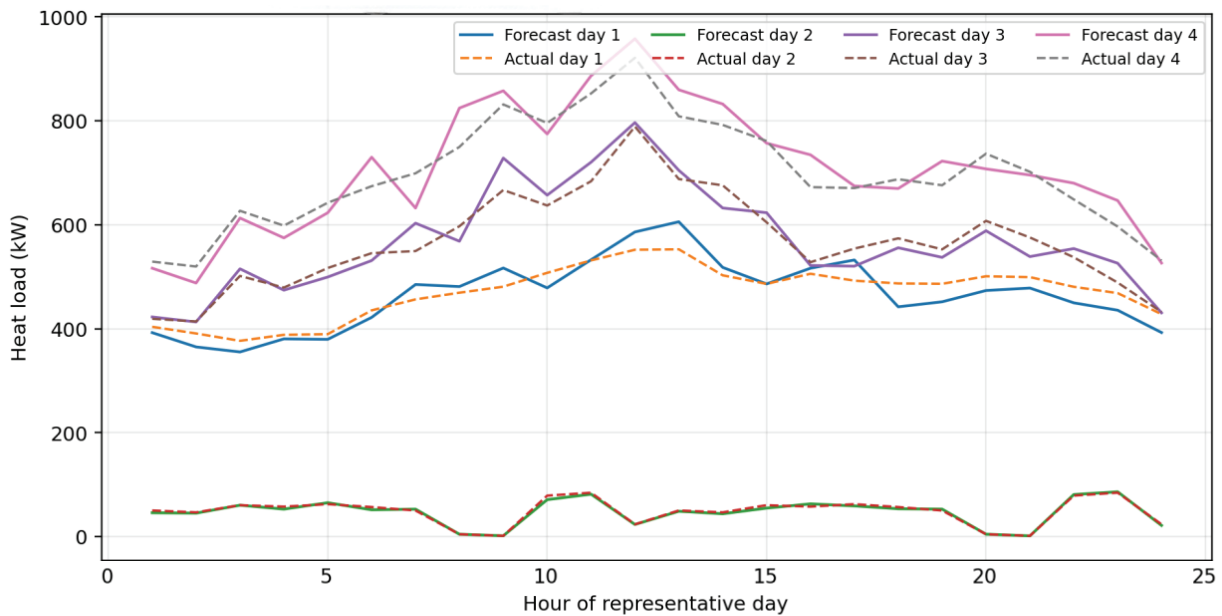
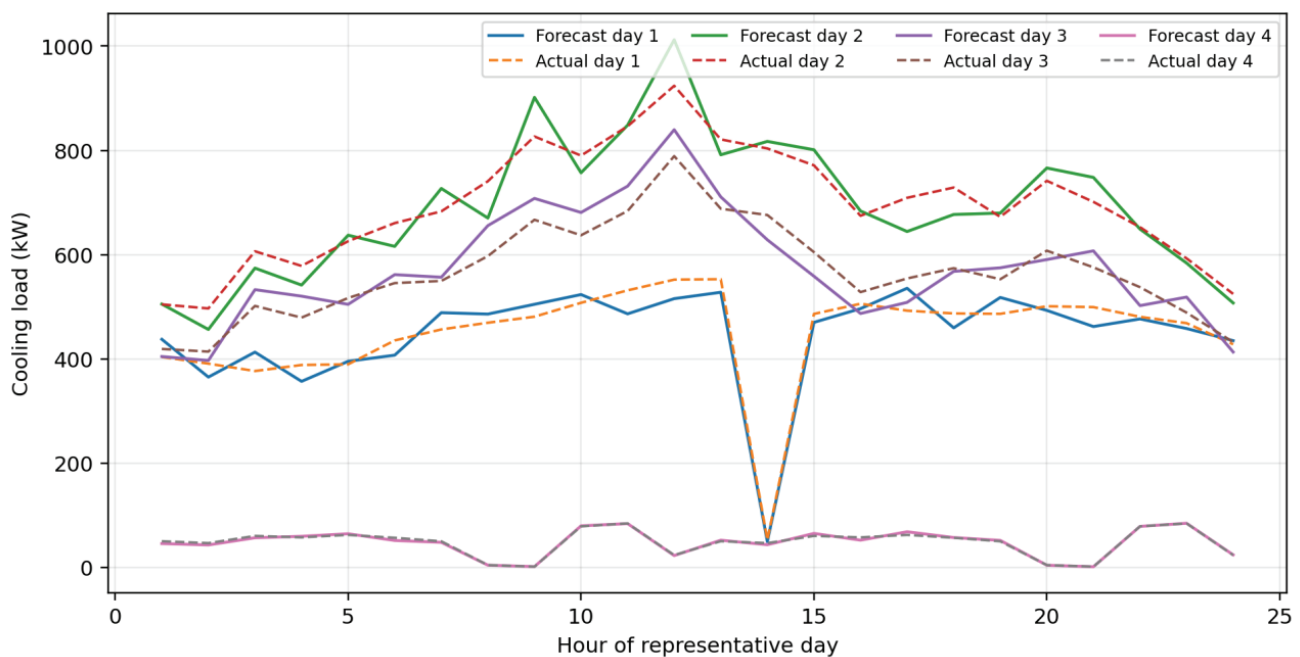


Figure 4: Forecast and realized cooling loads.

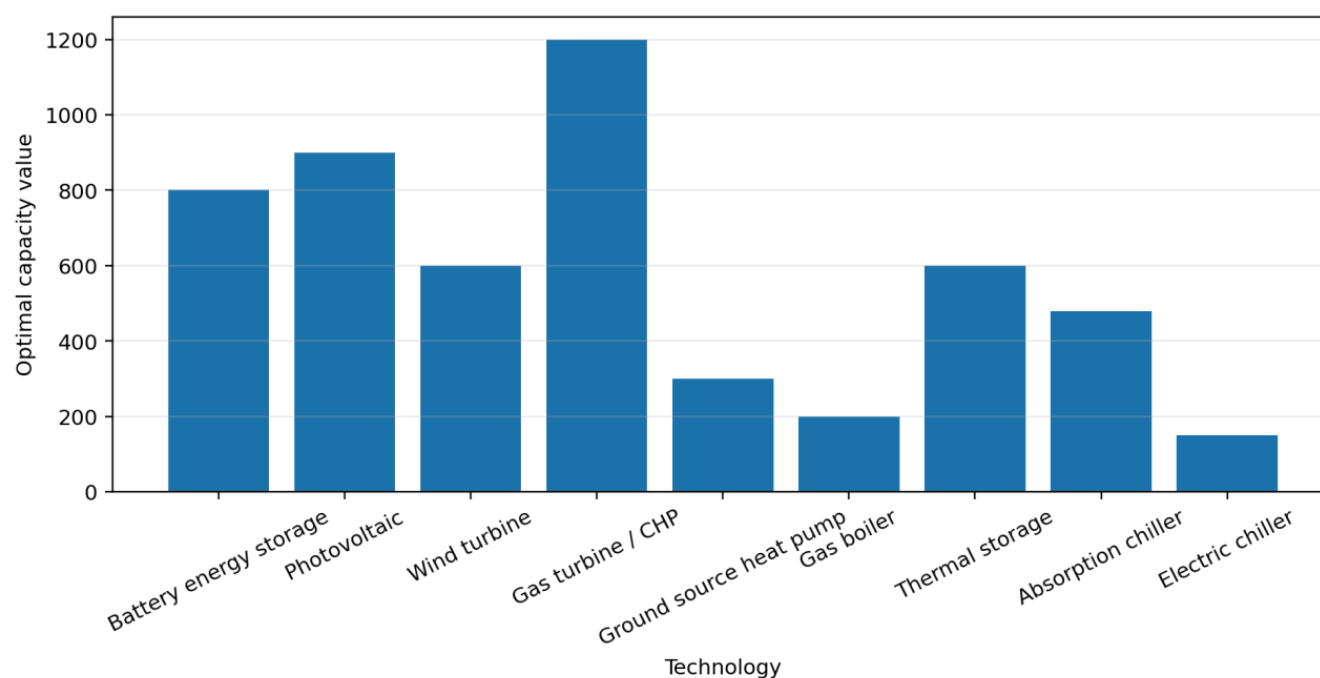


6. Results and Discussion

6.1 Optimal Capacity Configuration

The optimal capacity mix obtained from the benchmark model is reported in Table 2 and illustrated in Figure 5. Several structural features are noteworthy. First, the capacities of photovoltaic generation, wind generation, battery storage, the gas turbine, thermal storage, and the electric chiller all reach the upper bounds of their admissible planning intervals. This result indicates that, under the present demand patterns and price structure, renewable penetration and multi-timescale flexibility provide substantial system value. Second, the ground source heat pump is selected at its lower bound, suggesting that heat supply from CHP and auxiliary heating already satisfies most thermal requirements cost-effectively. Third, the absorption chiller converges to an interior optimum of 479.08 kW rather than to a boundary solution, which implies that the cooling structure is determined by a nontrivial temporal trade-off between electricity-driven and heat-driven refrigeration.

Figure 5: Optimal capacity configuration of the integrated energy system.



The daily equivalent annualized investment cost is 6286.79. Among all technologies, the CHP unit contributes the largest share to the investment term, followed by photovoltaic and wind generation. This ranking is consistent with the relatively high unit capital cost of the gas turbine and the large renewable capacities selected by the optimization.

6.2 Representative-Day Dispatch Characteristics

Table 5 summarizes the main load and supply quantities on each representative day. Representative days 2 and 4 are the most electricity-intensive, while representative day 4 also exhibits the highest heat demand. Representative day 2 is dominated by cooling demand, which activates both the electric chiller and the absorption chiller. By contrast, representative day 4 is primarily heat-driven and has only limited cooling demand.

Table 5: Representative-day load and supply summary.

Day	Ele. load	Heat load	Cool load	WT	PV	Grid buy	Grid sell
1	11028.10	11160.51	10763.60	9648.71	3443.96	1744.54	5409.76
2	16452.60	1130.65	16596.38	9666.06	5203.46	1939.58	4040.73
3	13429.73	13665.97	13761.85	8383.59	3414.35	3515.05	3503.67
4	16700.50	16988.81	1150.91	8514.85	2566.94	4422.79	2363.67

The dispatch profiles in Figures 6–8 reveal the internal coordination mechanism of the system. On representative day 2, strong wind and photovoltaic generation reduce net grid dependence during several hours, while the electric chiller and the ground source heat pump increase electrical demand in response to the high cooling load. At the same time, the CHP unit supplies recoverable heat that can be converted into cooling by the absorption chiller, thereby partially shifting the cooling burden away from the electrical side. On representative day 4, the thermal subsystem becomes the dominant operational focus: CHP heat, the ground source heat pump, the gas boiler, and thermal storage jointly satisfy the high heating requirement. These results demonstrate that the value of an integrated energy system does not arise merely from technology coexistence, but from the ability to redirect energy between carriers according to hourly marginal conditions.

Figure 6: Representative day 2 electric dispatch profile.

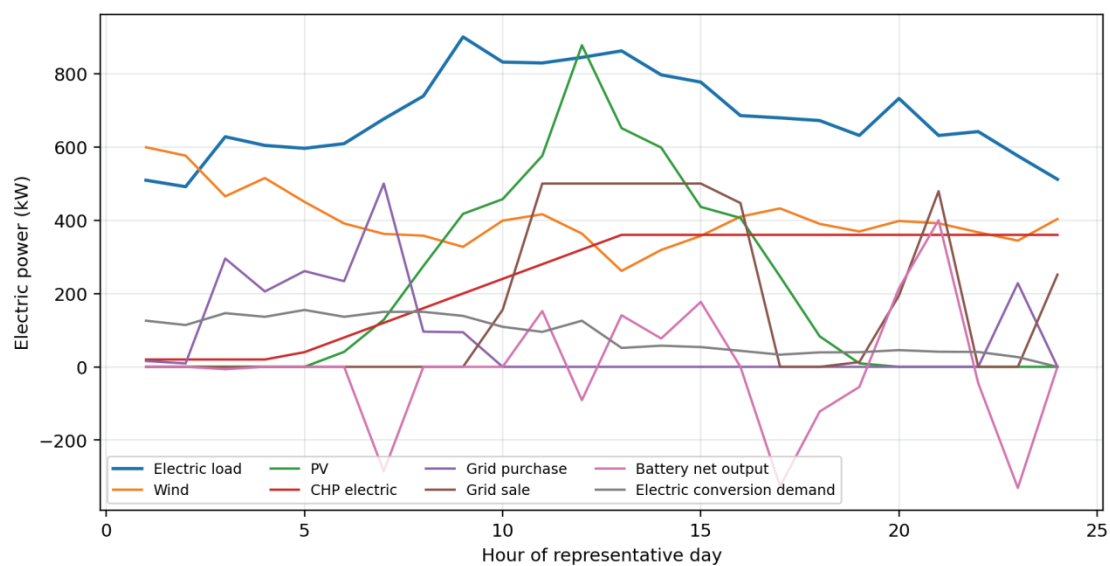


Figure 7: Representative day 4 heat dispatch profile.

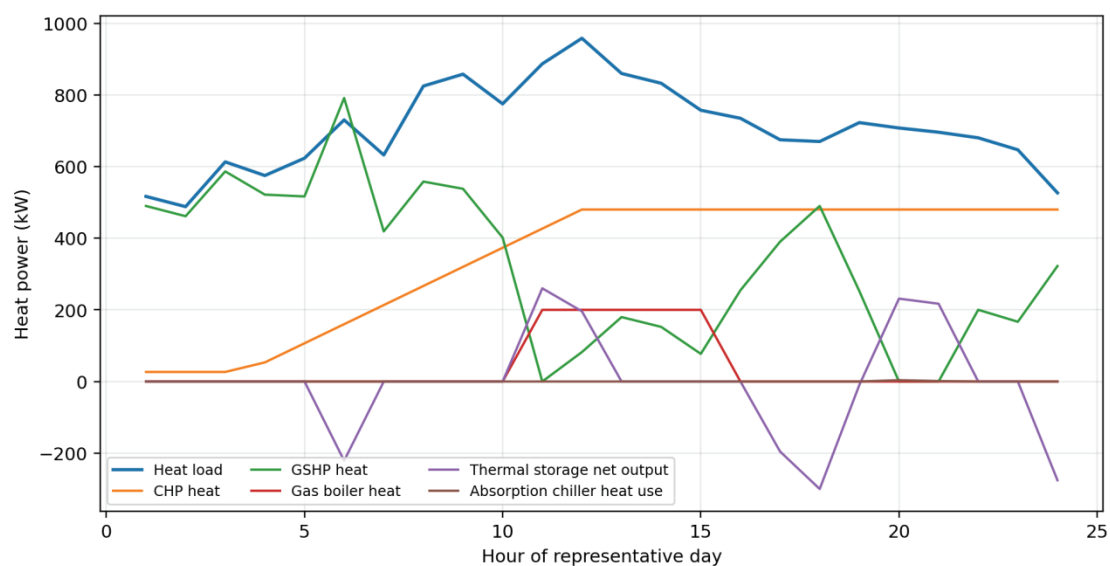
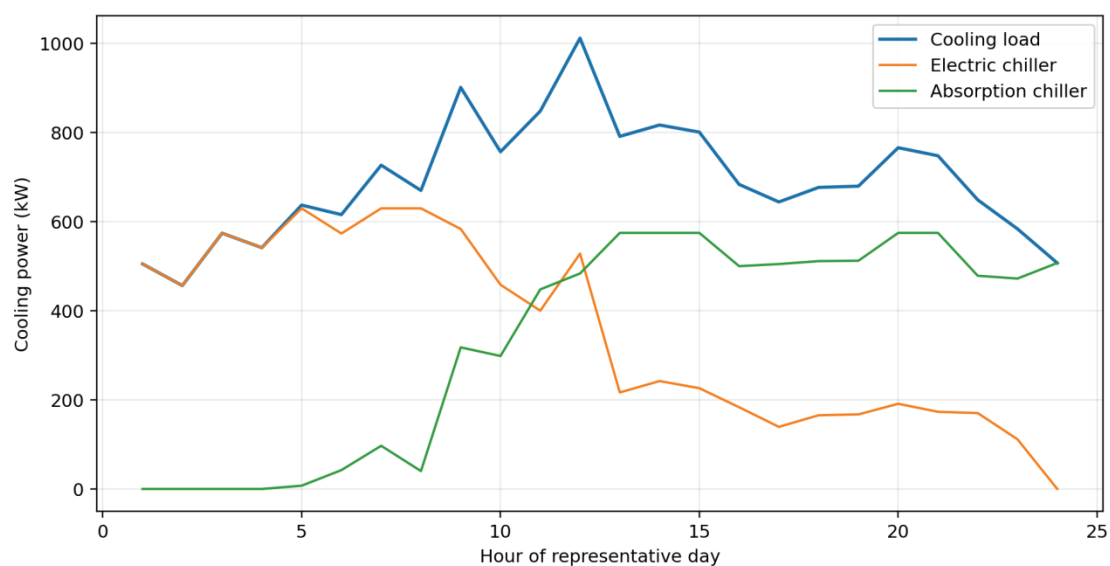


Figure 8: Representative day 2 cooling dispatch profile.



6.3 Economic and Ex-Post Operating Performance

Table 6 reports the forecast and realized operating results of each representative day. The total day-ahead composite operating objective over the four representative days is 77400.43, whereas the ex-post composite operating objective under realized loads is 78763.31. The increase is moderate, indicating that the selected capacity mix and dispatch structure remain robust under forecast deviation. The complete planning–operation objective, including the daily equivalent annualized investment term, is 83687.21.

Table 6: Representative-day operating results.

Day	F. direct	F. emission	F. comp.	A. direct	A. emission	A. comp.
1	11294.67	2024.31	13318.98	11615.95	2379.45	13995.39
2	14125.36	4159.39	18284.75	14383.25	4418.56	18801.81
3	14225.49	7028.15	21253.64	14410.05	7192.17	21602.22
4	15444.79	9098.28	24543.07	15450.34	8913.54	24363.88

Several insights can be drawn from these results. First, available renewable power is effectively utilized because it directly reduces grid purchases and weakens the environmental penalty associated with external electricity supply. Second, the CHP unit acts as a structural anchor of the system by simultaneously supporting the electricity and heat balances and by enabling absorption cooling through recoverable heat. Third, the joint presence of battery storage and thermal storage improves temporal flexibility on both the electrical and thermal sides, allowing the system to shift energy across hours and to mitigate part of the mismatch between renewable availability and demand. Fourth, the coexistence of the electric chiller and the absorption chiller allows the cooling load to be allocated dynamically between electricity and heat, which is particularly valuable when renewable output, CHP heat availability, and grid prices vary significantly across the day.

From an ex-post perspective, the gap between the forecast and realized objectives is mainly attributable to additional electricity purchase and compensatory thermal production when realized load exceeds the forecast level. Nevertheless, the deviation remains limited, which suggests that the optimized capacity portfolio provides adequate flexibility for representative forecast error levels.

7. Conclusion

This paper presents a complete reformulation of a bilevel integrated energy system model for coordinated capacity planning and day-ahead scheduling. The upper level determines the capacities of renewable generation, conversion devices, and storage technologies, while the lower level coordinates electricity, heat, cooling, and gas flows over representative days. The model explicitly represents renewable generation, CHP coupling, heating and cooling conversion, bidirectional grid interaction, storage dynamics, and gas consumption, and it is supported by a unified notation system and a detailed parameter specification.

To improve both interpretability and reproducibility, the study retains the original GA–PSO planning philosophy and further establishes an equivalent mixed-integer linear benchmark model. The numerical results show that the optimal configuration favors strong renewable deployment together with relatively large CHP, battery, and thermal-storage capacities. The representative-day dispatch results further demonstrate that coordinated use of CHP, heat pumping, electricity-driven cooling, and heat-driven cooling can effectively manage the multi-carrier demand structure. Overall, the study confirms that bilevel planning with explicit electricity–heat–cooling coupling provides an effective framework for the design and operation of integrated energy systems with renewable uncertainty and interdependent conversion technologies.

Funding

No

Conflict of Interests

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

Reference

- [1] M. Geidl, G. Koepfel, P. Favre-Perrod, B. Kloeckl, G. Andersson, and K. Froehlich, "Energy hubs for the future," *IEEE Power & Energy Magazine*, vol. 5, no. 1, pp. 24–30, 2007.
- [2] M. Geidl and G. Andersson, "Optimal power flow of multiple energy carriers," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 145–155, 2007.
- [3] G. Chicco and P. Mancarella, "Distributed multi-generation: A comprehensive view," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 3, pp. 535–551, 2009.
- [4] C. Marnay, G. Venkataramanan, M. Stadler, A. S. Siddiqui, R. Firestone, and B. Chandran, "Optimal technology selection and operation of commercial-building microgrids," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 975–982, 2008.
- [5] E. Fabrizio, V. Corrado, and M. Filippi, "A model to design and optimize multi-energy systems in buildings at the design concept stage," *Renewable Energy*, vol. 35, no. 3, pp. 644–655, 2010.
- [6] J. Keirstead, M. Jennings, and A. Sivakumar, "A review of urban energy system models: Approaches, challenges and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3847–3866, 2012.
- [7] S. Pfenniger, A. Hawkes, and J. Keirstead, "Energy systems modeling for twenty-first century energy challenges," *Renewable and Sustainable Energy Reviews*, vol. 33, pp. 74–86, 2014.
- [8] P. Mancarella, "MES (multi-energy systems): An overview of concepts and evaluation models," *Energy*, vol. 65, pp. 1–17, 2014.
- [9] J. Wu, J. Yan, H. Jia, N. Hatzigiorgiou, N. Djilali, and H. Sun, "Integrated energy systems," *Applied Energy*, vol. 167, pp. 155–157, 2016.
- [10] J. Reynolds, M. Ahmad, and Y. Rezgui, "Holistic modelling techniques for the operational optimization of multi-vector energy systems," *Energy and Buildings*, vol. 169, pp. 397–416, 2018.
- [11] R. Rigo-Mariani, S. O. C. Wae, S. Mazzoni, and A. Romagnoli, "Comparison of optimization frameworks for the design of a multi-energy microgrid," *Applied Energy*, vol. 257, p. 113982, 2020.
- [12] S. A. Mansouri, A. Ahmarinejad, M. Ansarian, M. S. Javadi, and J. P. S. Catalao, "Stochastic planning and operation of energy hubs considering demand response programs using Benders decomposition approach," *International Journal of Electrical Power & Energy Systems*, vol. 120, p. 106030, 2020.
- [13] A. E. H. Berjawi, S. L. Walker, C. Patsios, and S. H. R. Hosseini, "An evaluation framework for future integrated energy systems: A whole energy systems approach," *Renewable and Sustainable Energy Reviews*, vol. 145, p. 111163, 2021.
- [14] M. H. Shams, M. Shahabi, M. MansourLakouraj, M. Shafie-khah, and J. P. S. Catalao, "Adjustable robust optimization approach for two-stage operation of energy hub-based microgrids," *Energy*, vol. 222, p. 119894, 2021.
- [15] P. C. Taylor, M. Abeysekera, Y. Bian, D. Cetanovic, M. Deakin, A. Ehsan, V. Levi, F. Li, R. Oduro, R. Preece, P. G. Taylor, V. Terzija, S. L. Walker, and J. Wu, "An interdisciplinary research perspective on the future of multi-vector energy networks," *International Journal of Electrical Power & Energy Systems*, vol. 135, p. 107492, 2022.
- [16] T. Ding, W. Jia, M. Shahidehpour, O. Han, Y. Sun, and Z. Zhang, "Review of optimization methods for energy hub planning, operation, trading, and control," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 3, pp. 1802–1818, 2022.
- [17] Y. Zhou, X. Li, H. Han, Z. Wei, H. Zang, G. Sun, and S. Chen, "Resilience-oriented planning of integrated electricity and heat systems: A stochastic distributionally robust optimization approach," *Applied Energy*, vol. 353, p. 122053, 2024.
- [18] X. Wu, B. Zhang, M. P. Nielsen, and Z. Chen, "Multi-stage planning of integrated electricity-gas-heating system in the context of carbon emission reduction," *Applied Energy*, vol. 358, p. 122584, 2024.
- [19] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI, USA: University of Michigan Press, 1975.
- [20] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, 1995, pp. 1942–1948.
- [21] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004, pp. 284–289.
- [22] Gurobi Optimization, LLC, *Gurobi Optimizer Reference Manual*. Beaverton, OR, USA, 2024.

Research and Development Trends of Expressway Maintenance Management Technology

Qingsong Zhang*

He'an Expressway Management Center, Anhui Transportation Holding Group Co., Ltd., Anqing, China

*Corresponding author: Qingsong Zhang, qsz9805@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: Expressway safety is a crucial prerequisite for safeguarding the life and health of drivers and passengers, with maintenance management serving as its core support. To enhance the safety and sustainability of expressway maintenance management, this paper first analyzes the fundamental components, characteristics, existing problems, and corresponding countermeasures of maintenance management. Secondly, it proposes optimization measures for the safety management of expressway maintenance construction. Finally, it predicts future development trends in maintenance operations from four perspectives: management mechanisms, digital transformation, investment mechanism optimization, and full life-cycle ecologization. This study aims to provide decision-making support and technical references for expressway operating entities and maintenance personnel.

Keywords: Expressway; Maintenance Management; Construction Safety; Status Analysis

Published: Apr 5, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1235>

1. Introduction

Since the opening of mainland China's first expressway (the Shanghai–Jiading Expressway) in 1988, China's expressway construction has achieved leapfrog development ^[1]. With socioeconomic progress, the continuous increase in the number of vehicles has directly led to a substantial rise in expressway traffic volume ^[2]. Under long-term load effects, pavement structures have experienced varying degrees of damage ^[3], while some traffic facilities are also facing aging and deterioration, making maintenance and rehabilitation work increasingly challenging. The long-standing tendency of “prioritizing construction over maintenance” has resulted in insufficient investment in maintenance infrastructure and relatively outdated management practices ^[4]. In actual operation, maintenance activities not only affect traffic efficiency but also cause secondary accidents due to inadequate safety protection, threatening the safety of the public and constraining the sustainable development of the expressway network.

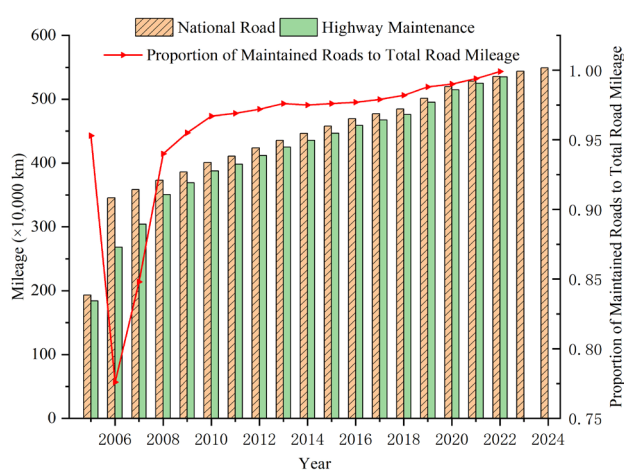
With the rapid advancement of intelligent driver assistance technologies, safety management in expressway maintenance work zones is facing new challenges. Establishing a modern maintenance construction system that is “safe, efficient, and intelligent” has become a key focus of industry development ^[5]. Focusing on the critical issue of safety in expressway maintenance work zones, this paper systematically reviews the current state of expressway maintenance management and construction safety, analyzes the safety hazards existing in current maintenance management, and explores future development directions, aiming to provide a reference for improving expressway maintenance management.

2. Analysis of Current Expressway Maintenance Status

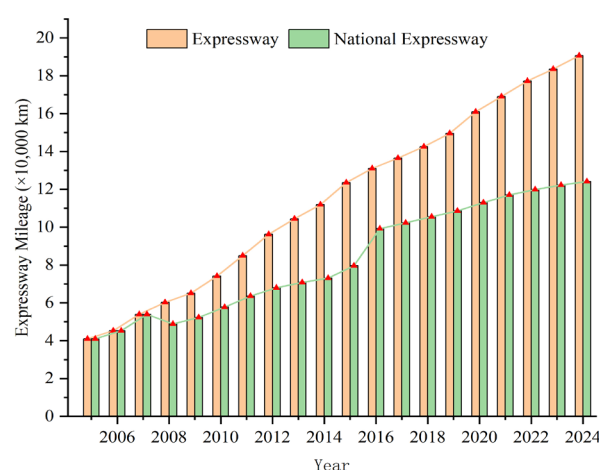
By the end of 2024, the total road mileage in China reached 5,940,400 kilometers, an increase of 53,600 kilometers over the previous year. The expressway mileage stood at 190,700 kilometers, a year-on-year increase of 7,000 kilometers, of which national expressways accounted for 124,100 kilometers, an increase of 1,800 kilometers year-on-year. According to data from the Ministry of Transport (Figure 1), both total road mileage and expressway mileage in China exhibited a sustained growth trend from 2005 to 2024. Among these, the total road mileage grew most rapidly in 2006 (a year-on-year increase of 79.07%), primarily due to a statistical adjustment that included village roads in the national total road mileage for the first time. Expressway mileage reached its peak growth rate in 2016 (a year-on-year increase of 24.62%), closely related to the network densification policies during the 13th Five-Year Plan period. Although road maintenance mileage increased year by year from 2005 to 2022 (data for 2023–2024 have not yet been released), its proportion relative to total road mileage dropped sharply from 95.3% to 77.60% between 2005 and 2006. This contrast stems from a combination of statistical and policy factors: maintenance data for 2005 and earlier covered only trunk roads such as national highways, provincial highways, and county roads. However, after village roads were included in the total mileage in 2006, rural roads—constrained by issues such as funding shortages and unclear responsibility entities—maintained a consistently lower maintenance coverage rate than trunk roads, leading to a decline in the overall proportion. National expressway mileage experienced a phased decline from 2007 to 2008 (from 53,900 kilometers in 2007 to 48,900 kilometers in 2008), primarily due to the refinement of statistical classification. Data for 2007 and earlier included some provincial-level expressways. Following the implementation of the National Expressway Network Plan in 2008, a strict distinction was made between the two categories, and some routes originally classified as “national expressways” were reclassified as provincial expressways, resulting in a downward adjustment of national expressway mileage. Thereafter, with the advancement of network construction, national expressway mileage resumed its growth trend, reaching 124,100 kilometers by 2024.

Figure 1. National road and expressway mileage

a. National Road Mileage and Maintenance Proportion



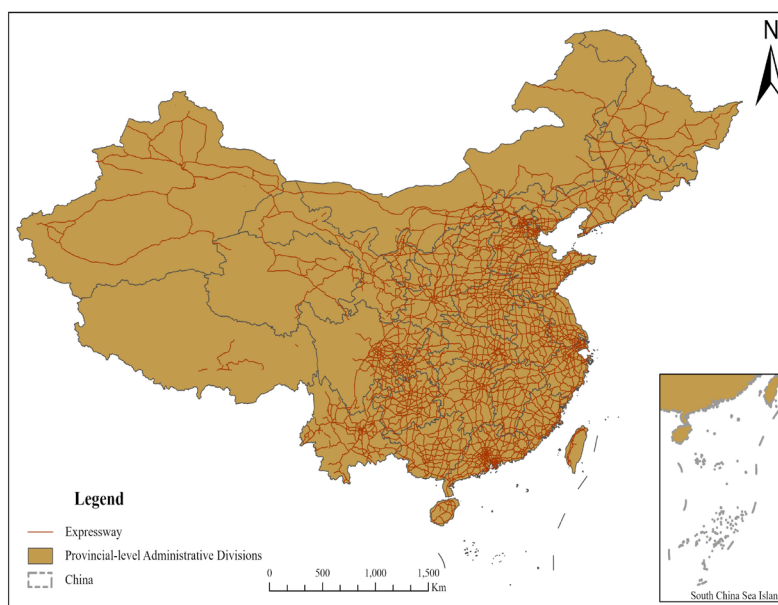
b. National Expressway Mileage



Source: Ministry of Transport of the People's Republic of China. (<https://www.mot.gov.cn/>)

This study is based on 2024 expressway vector data of China from OpenStreetMap (<https://openstreetmap.org>) and Chinese administrative division data released by the National Catalogue of Service for Geographic Information Resources (<https://www.webmap.cn/>). Spatial registration and overlay analysis were conducted using the ArcGIS Pro platform. As shown in Figure 2, the spatial distribution of expressways in China exhibits significant regional imbalance: the eastern region features a dense road network with high coverage; while the western region has longer individual expressway routes (e.g., the Lianyungang–Khorog Expressway), the number of routes and network density are considerably lower than those in the eastern region. This pattern closely aligns with China's topographic gradients and the distribution of population and economic activity^[6].

Figure 2. National expressway distribution in China



With the rapid expansion of China's expressway network, the importance of maintenance management has become increasingly prominent. Compared with ordinary roads, expressways are characterized by complex road conditions, high traffic volumes, and high vehicle speeds, resulting in significantly higher rates of pavement wear and facility aging. Consequently, stricter requirements are imposed on the timeliness, technical standards, and safety of maintenance management. To systematically analyze the core challenges and practical experiences of current expressway maintenance management, this paper integrates and reviews key aspects of the maintenance management system through policy document analysis, technical specification comparison, and case studies of typical road sections^[7,8].

Expressway maintenance possesses numerous characteristics. This paper provides a detailed introduction from the following aspects.

(1) Dispersion: For expressways in operation, factors such as high traffic volume, high vehicle speeds, and natural disasters result in the occurrence of pavement distress exhibiting significant randomness and spatial discreteness. This dispersion is primarily reflected in two aspects: the distribution of distress and the organization of maintenance operations. In maintenance operations, routine tasks such as median strip garbage removal, replacement of anti-glare panel reflective films, and cleaning of longitudinal drainage ditches are often not coordinated due to the dispersed timing of operations, leading to multiple lane closures. Furthermore, due to the division of professional responsibilities among different functional departments, repair work on the same road section often requires repeated lane closures as a result of uncoordinated construction schedules. This phenomenon of repeated lane closures not only leads to inefficient allocation of maintenance resources but also causes additional travel inconvenience for drivers and passengers.

(2) Service Orientation: As a leading basic industry of the national economy, expressways possess significant public welfare attributes, and their operation and management are closely related to public life. In the process of maintenance management, not only economic benefits but also social benefits must be considered. Particularly during peak traffic periods such as holidays, failure to properly coordinate the balance between maintenance operations and traffic demand may lead to severe traffic congestion and social impact. In addition, heavy rainfall during the rainy season can easily trigger slope failures and drainage ditch collapses, which not only affect the drainage function of the road but may also endanger adjacent farmland and residential facilities. Therefore, maintenance operations must strictly comply with the regulations of traffic management departments, scientifically formulate construction plans, and efficiently complete emergency repair tasks under the premise of ensuring traffic safety.

(3) High Risk: Compared with ordinary roads, expressways exhibit more complex road conditions, including not only special structures such as tunnels, river-crossing bridges, and high slopes but also higher vehicle speeds. This particularity imposes

higher requirements on maintenance operations: on the one hand, vehicles traveling through maintenance work zones constitute an uncontrollable risk source; on the other hand, various uncertain factors such as weather changes and equipment failures during construction collectively form a complex safety risk system.

(4) Complexity: The complexity of expressway maintenance management is primarily reflected in multidimensional systemic challenges. In terms of road structure, the coexistence of differentiated structures such as tunnel groups, river-crossing bridges, high slopes, and interchange hubs requires maintenance operations to simultaneously possess multidisciplinary expertise in bridge engineering, tunnel engineering, geotechnical engineering, and other fields, while different structures exhibit significant differences in maintenance cycles and technical standards. Regarding dynamic risks, the average vehicle speed of 100 km/h necessitates dynamic calculation of the buffer distance in work zones, alongside the establishment of a 15-minute rapid response mechanism for natural disasters such as heavy rainfall and landslides. Furthermore, multi-department coordination is also an important feature, requiring the integration of permit approvals from road administration departments, traffic organization from traffic police departments, and process coordination from construction units. By establishing a “multi-party coordination” joint service mechanism, traffic safety and operational efficiency in construction zones can be ensured.

Expressway maintenance work is primarily divided into two categories: “routine maintenance” and “maintenance engineering.” Maintenance engineering further includes four subcategories: preventive maintenance, remedial maintenance, special maintenance, and emergency maintenance, as detailed in Table 1. Regardless of the type, maintenance is conducted to prevent and address issues related to seven key aspects: subgrade, pavement, bridges, tunnels, traffic safety facilities, electromechanical facilities, management and service facilities, as well as landscaping and environmental protection facilities.

Table 1. Classification of highway maintenance engineering operations

Expressway maintenance	Maintenance Category	Description
Routine Maintenance	—	Subgrade
		Pavement
		Bridges
		Tunnels
		Traffic Safety Facilities
		Electromechanical Facilities
		Management and Service Facilities
		Landscaping and Environmental Protection Facilities
Maintenance Engineering	Preventive Maintenance	Subgrade
		Pavement
		Bridges and culverts
		Tunnels
	Remedial Maintenance	Subgrade
		Pavement
		Bridges and culverts
		Tunnels
		Traffic Safety Facilities
		Electromechanical Facilities
		Management and Service Facilities
		Landscaping and Environmental Protection Facilities
	Special Maintenance	Various Facilities
	Emergency Maintenance	Various Facilities

Despite the continuous increase in government investment and technical support for maintenance management, current expressway maintenance still faces numerous challenges. First, the identification of potential risks remains predominantly reliant on traditional manual methods, which are not only time-consuming and labor-intensive but also struggle to achieve comprehensive coverage of high-risk areas such as steep geological regions. Second, the maintenance philosophy lags behind, with a prevailing reactive approach of “addressing issues as they arise,” lacking systematic and forward-looking integrated governance strategies. Third, maintenance methods urgently require upgrading. Although intelligent maintenance technologies have been piloted in some operating entities, front-line maintenance personnel still primarily rely on foot patrols, resulting in low management efficiency. Furthermore, the multi-department coordination mechanism is inadequate. Due to unclear division of management responsibilities, the same road section is often subject to duplicate inspections, leading to resource waste. Finally, maintenance resources are particularly scarce in complex sections such as mountainous areas. The maintenance of special structures like tunnels, bridges, and high slopes requires multidisciplinary expertise; however, constrained by industry compensation levels, there is a severe shortage of highly qualified maintenance personnel.

3. Analysis of Safety Management in Expressway Maintenance Construction

With the end of the traditional “prioritizing construction over maintenance” model, the highway maintenance industry is facing unprecedented development opportunities. This transformation has prompted a large number of construction enterprises to expand into the maintenance sector. However, the rapid increase in market participants has also brought issues such as lowered industry entry barriers and inconsistent implementation of construction standards. Consequently, maintenance operations are confronted with multiple safety challenges. In practice, factors such as improper placement of construction signs, insufficient safety personnel, weak safety awareness among workers, and the uncontrollable nature of high-speed passing vehicles have led to frequent safety accidents in maintenance work zones^[9]. In response to the above issues, China has established clear provisions in maintenance engineering technical standards, formulating a series of detailed specifications and guidelines. However, the effectiveness of their implementation is constrained by the uneven quality of practitioners^[10]. Empirical studies indicate that construction teams receiving standardized training can significantly reduce the incidence of operational accidents. Therefore, it is necessary to establish a tiered training and certification mechanism.

As a core component of the modern transportation infrastructure operation and maintenance system, expressway maintenance construction holds strategic value in two dimensions: at the physical level, it ensures the structural integrity and functional continuity of the road through systematic maintenance; at the socioeconomic level, it directly affects public travel safety and regional logistics efficiency. To maximize maintenance effectiveness, it is necessary to construct a “Quality-Efficiency-Safety” (QES) collaborative optimization model based on full life-cycle management^[11]. Specifically, this system comprises two key elements: first, a scientific planning system based on pavement performance deterioration laws, integrating multi-source data-driven distress diagnosis (using dual assessment of PCI and IRI indicators), dynamic resource allocation algorithms, weather window optimization models, and tiered emergency response plans. Second, a standardized management system relies on the PDCA cycle mechanism, clarifies the responsibilities of all participants through the RACI responsibility matrix, and leverages a BIM+GIS collaborative platform to achieve real-time interaction of design parameters, construction progress, and quality inspection data^[12].

Safety management in expressway maintenance construction is a systematic endeavor requiring multi-department coordination. Before commencing maintenance operations, construction units must strictly adhere to the “Three Advances” principle: submit detailed traffic organization plans to traffic police, road administration departments, and expressway operating entities 72 hours in advance for joint review; conduct comprehensive hazard inspections of the work area using technological means such as UAV aerial photography and ground-penetrating radar in advance, ensuring inspection coverage of no less than 95%; and conduct quarterly emergency drills for high-risk scenarios such as slope collapses and vehicle intrusions in advance, ensuring emergency response times are controlled within 15 minutes.

In terms of management measures, maintenance construction management departments should focus on three key areas: first, implement differentiated control measures in accordance with the requirements of the Safety Work Rules for Highway Maintenance (JTG H30), developing customized “one-road-one-policy” plans for special sections such as long steep slopes

and tunnel groups. Second, strengthen intelligent supervision by promoting the application of vehicle-mounted video surveillance systems equipped with AI recognition capabilities to monitor critical aspects such as safety equipment wearing and construction sign placement in real time. The system currently applied in Jiangsu Province has achieved a recognition accuracy of 92%. Third, enforce strict qualification management by establishing an “one-person-one-file” electronic certificate database and achieving real-time integration with the National Highway Construction Market Credit Information System, thoroughly eliminating irregular practices such as certificate leasing.

The construction of the personnel training system should focus on three priorities: implement a three-tier safety education system, requiring new employees to complete 24 hours of online theoretical training and 8 hours of VR accident simulation drills; strictly enforce the requirement of working with certificates, requiring special operation personnel to obtain the Highway Maintenance Operation Post Certificate issued by the Ministry of Transport; and adhere to case-based safety education, organizing monthly analyses of typical accident cases, such as the profound lesson from the March 29, 2025, incident in Tongling, Anhui, where a new energy vehicle collided with a concrete barrier and caught fire.

In terms of technological innovation, active promotion of mechanization and intelligent transformation should be pursued: advanced equipment such as laser obstacle removal devices should be adopted to replace manual operations, achieving operational efficiency up to six times that of traditional methods. Digital twin technology should be promoted to simulate construction traffic impacts based on BIM models and optimize traffic diversion plans. A project in Zhejiang Province achieved a 37% reduction in congestion complaints on construction sections after application. Additionally, a sound construction safety credit evaluation system should be established, implementing a blacklist system for violations to form a strong deterrent. Through these multiple measures, a comprehensive, multi-dimensional safety assurance system for maintenance construction can be established.

4. Future Development Trends of Expressway Maintenance Operations

4.1 Management Mechanism Innovation

The conceptual limitations of expressway maintenance managers are primarily reflected in the failure to carry out maintenance work from a full life-cycle perspective, resulting in a disconnect between the construction and operation phases. Operating entities typically take over maintenance only after the road opens to traffic, making it difficult for them to fully grasp the structural information of the road sections or develop customized maintenance plans (such as “one-road-one-policy”). To address management gaps, additional human and material resources are often required. Future efforts should focus on management mechanism innovation to resolve this issue, including: organizing regular management training, facilitating cross-provincial exchanges on real-time road conditions and maintenance technologies, formulating a unified national Smart Maintenance Technical Standard (clarifying specifications for UAV inspections, IoT-based detection, etc.), and developing an integrated system platform to enable full-process online management of defect reporting, task assignment, and acceptance evaluation ^[13].

4.2 Digital Transformation

With the rapid advancement of artificial intelligence technology, many industries have successfully leveraged AI to enhance work efficiency. However, some expressway operating entities still primarily rely on traditional manual inspection methods for maintenance operations. This approach, constrained by the professional competence of maintenance personnel, often fails to detect hidden distress such as subgrade cracks and internal structural damage. To improve maintenance effectiveness, the future should see vigorous promotion of intelligent monitoring technologies: automated identification of surface distress such as pavement cracks and subsidence through “UAV + AI image recognition”; real-time monitoring of structural conditions such as bridge stress and tunnel deformation via embedded fiber optic sensor networks; development of pavement performance deterioration prediction models based on big data analytics to scientifically determine optimal maintenance timing; and simulation-based comparison of the economic viability of different maintenance alternatives using BIM-GIS integrated platforms. The coordinated application of these intelligent methods will drive expressway maintenance from reactive repair to proactive prevention, significantly enhancing maintenance efficiency while ensuring traffic safety ^[14].

4.3 Investment Mechanism Optimization

Although the state has continuously increased its emphasis on highway maintenance, the compensation for maintenance personnel remains relatively low, resulting in a sustained loss of multidisciplinary talent. To improve this situation, the following measures are recommended: first, reasonably increase the income levels of maintenance management personnel to attract professionals with expertise in bridges, tunnels, and pavement, thereby alleviating the current shortage of specialized talent^[15]; second, organize regular professional skills training to continuously enhance the competence of the maintenance workforce; at the same time, establish a special fund for geological disaster prevention and control on mountainous expressways to ensure dedicated use of funds—a particularly necessary measure given the complex road conditions in mountainous areas. Additionally, special incentives should be provided to individuals who achieve results in maintenance process innovation and key technical research. Finally, based on the maintenance needs of specific road sections, the application of new materials, new processes, and new technologies should be actively promoted to reduce safety risks in maintenance operations.

4.4 Full Life-Cycle Ecologization

Based on the current state of expressway development in China, the concept of full life-cycle ecologization should be comprehensively promoted. Specific measures include: first, promoting 100% recycling technology for asphalt pavement (for example, Shandong Province has achieved a recycling rate of 92% for old pavement materials), thereby improving the level of resource circular utilization; second, constructing integrated “photovoltaic-storage-charging” energy stations in service areas and tunnels to supply power for facilities and reduce electricity waste; third, applying bioengineering techniques for slope stabilization to prevent water-induced slope damage (such as the demonstration technology of vine plant slope stabilization applied in Yunnan Province); at the same time, planting vegetation with dust absorption and noise reduction functions in the central median to further enhance ecological benefits. Through multi-dimensional optimization, coordinated development of expressway maintenance and ecological environment protection can be achieved^[16].

5. Conclusion

This paper elaborates on the current basic content of maintenance management from three aspects: the current status of expressway maintenance, safety management in maintenance construction, and future development trends of maintenance operations, providing reference suggestions for maintenance personnel to systematically understand the fundamental aspects of maintenance work. Expressway construction and maintenance are important links in ensuring the normal use of expressways. Preliminary construction encompasses the entire process from planning to opening, while subsequent maintenance provides sustainability assurance. The two complement each other; through construction-maintenance synergy, the level of expressway safety protection can be jointly enhanced, ensuring efficient travel for drivers and passengers. Expressway maintenance management must adhere to the concept of sustainable development, taking a holistic approach by fully considering the needs of later maintenance during the initial construction phase, and integrating preventive maintenance concepts into design, construction, and other stages.

To comprehensively improve maintenance levels, it is recommended to establish a modern, intelligent integrated expressway maintenance system: reasonably increase maintenance investment, establish dynamic adjustment mechanisms, introduce multidisciplinary talent, implement differentiated resource allocation, and improve the efficiency of fund utilization; leverage the coordination advantages of provincial operation platforms to establish cross-regional collaboration mechanisms through resource integration and unified management, achieving complementary experience and advantages; in terms of technological innovation, actively promote digital transformation, deeply apply cutting-edge technologies such as big data analytics, artificial intelligence decision-making, and smart highway systems, transforming the expressway into an integrated “management-maintenance-construction” comprehensive display system. This will not only effectively extend the service life of expressways but also promote the high-quality, sustainable development of the maintenance industry, providing solid support for the construction of a leading transportation nation.

Conclusion

In summary, LeBron James has demonstrated his unique value as an all-around basketball player through his outstanding performance in scoring, organization, defense and leadership. By comparing multiple key data of his career, including scoring efficiency, assists, defensive comprehensiveness and career durability, LeBron shows that he is not only a scoring machine, but also a core player who can lead the overall situation and drive the team. In particular, throughout his career, LeBron has been able to maintain a high level of performance and adapt to the needs of different teams by adjusting his role, which further proves his outstanding ability and flexibility in multiple positions. These traits have gradually established LeBron as a more complete basketball player in terms of data dimensions and overall comprehensiveness.

Although Michael Jordan is undoubtedly one of the most influential players in basketball history, his six championships and unparalleled competitive spirit have established his position as a basketball “myth”. LeBron has surpassed the traditional shooting guard role through his comprehensiveness and career continuity and has become an all-around symbol of contemporary basketball. Whether it is scoring, assists, or defense, LeBron has demonstrated top-notch abilities in all aspects, combined with his leadership performance, forming a unique basketball style. Readers can further think about who is the greatest player in basketball history and form their own opinions based on the data and analysis provided in this article.

Funding

No

Conflict of Interests

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

Reference

- [1] The birth of mainland China's first expressway. (2021). *Grand Garden of Science*, (17), 32–33.
- [2] Ke, W. (2015). Spatio-temporal characteristics of expressway traffic flow network and urban spatial interaction: A case study in Jiangsu Province [Doctoral dissertation]. Nanjing Normal University.
- [3] Wang, S. (2021). Investigation on fatigue cracking behavior of semi-flexible pavement [Doctoral dissertation]. Southeast University. <https://doi.org/10.27014/d.cnki.gdnau.2021.004844>
- [4] Zhang, W. (2012). Practice and reflection on the construction and maintenance of rural roads in mountainous areas. *Journal of Transport Management Institute Ministry of Transport of the People's Republic of China*, 22(1), 41–43.
- [5] Zhang, Y. (2019). Research on intelligent assisted driving control strategy based on traffic accident scene [Master's thesis]. Tianjin University of Technology and Education.
- [6] Wang, C., & Cheng, J. (2016). Spatial pattern of expressway network accessibility and evolution in China. *Geographical Science*, 36(6), 803–812. <https://doi.org/10.13249/j.cnki.sgs.2016.06.001>
- [7] Zhou, L. (2023). Analysis of main problems and measures in highway maintenance management. *Jiaotong Keji Yu Guanli*, 4(10), 174–176.
- [8] Interpretation of the “Highway Maintenance Technical Standards”. (2024). *Urban Roads Bridges & Flood Control*, (3), 273–274.
- [9] Li, W. (2025). Problems and solutions of safety production management of expressway maintenance. *Lecture Notes in Education, Arts, Management and Social Science*, 3(2), 176–182.
- [10] Chen, N. (2011). Study on construction peasant-labor training and appraisal system [Doctoral dissertation]. Tianjin University.
- [11] Massera Filho, C., Terra, M. H., & Wolf, D. F. (2017). Safe optimization of highway traffic with robust model predictive control-based cooperative adaptive cruise control. *IEEE Transactions on Intelligent Transportation Systems*, 18(11), 3193–3203. <https://doi.org/10.1109/TITS.2017.2679098>
- [12] Li, G. (2017). Research on quality management of S provincial highway life cycle [Master's thesis]. Shandong University.
- [13] Li, Q., Xi, M., & Liang, G. (2025). Analysis of the current situation and problems of expressway maintenance management in China. In *Proceedings of the 14th Academic Annual Conference of the Maintenance and Management*

- Branch of China Highway and Transportation Society (pp. 32–37). <https://doi.org/10.26914/c.cnkihy.2025.008685>
- [14] Wang, M. (2025). Research on the innovative management model of highway maintenance work zones empowered by digital technology. *China Construction Metal Structure*, 24(11), 167–169. <https://doi.org/10.20080/j.cnki.ISSN1671-3362.2025.11.057>
- [15] Guo, D., & Sun, Y. (2025). Current situation and countermeasure analysis of highway road maintenance construction management. *Auto Weekly*, (6), 51–53.
- [16] Bian, H. (2025). Analysis of the application of green construction technology in highway maintenance. *Theoretical Research in Urban Construction*, (15), 139–141. <https://doi.org/10.19569/j.cnki.cn119313/tu.202515047>

Research on the Construction of Digital Standard Contracts for Artificial Intelligence Data Training

Qiuhui Ren^{1*}, Xiaohua Fu²

1. School of Marxism, Henan Forestry Vocational College, Luoyang, Henan, 471002, China

2. School of Cultural Tourism, Chengdu Polytechnic, Chengdu, Sichuan, 610041, China

*Corresponding author: Qiuhui Ren, qiuhiuren0907@126.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 analysis and comparative study of domestic and foreign literature, this paper focuses on the legal regulatory issues arising from the process of artificial intelligence (AI) data training. It aims to address the inherent contradictions between machine learning models and the technical characteristics of traditional legal norms, the lack of algorithmic transparency, and the identification of data ownership by constructing digital standard contracts. Corresponding development suggestions are put forward from three aspects: establishing a dynamic update mechanism for contracts, standardizing transnational data training cooperation, and deepening the legal effect of contract execution results.

Keywords: Artificial Intelligence; Data Training Behavior; Digital Standard Contract

Published: Apr 12, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1246>

1. Introduction

1.1 Research Background

With the rapid development and wide application of AI technology, data training, as the core link of AI systems, involves the collection, processing and utilization of massive data, triggering a host of new legal issues in the legal field. Risks such as data privacy leakage, discriminatory results caused by algorithmic bias, and disputes over intellectual property ownership have become increasingly prominent, exposing the practical dilemma of the lack of effective legal regulation on AI data training behaviors^[1,2]. In addition, from intelligent judicial systems to automated administrative decision-making, AI has penetrated into all aspects of legal practice. While improving efficiency, this technological application has also brought about many legal challenges. There is no unified global standard for data training, and the legal conflicts among different jurisdictions and the limitations of industry self-regulatory norms have become increasingly prominent^[3]. These changes brought about by AI are triggering a profound transformation in the legal field. Against this background, it is particularly urgent to construct binding and universal digital standard contracts, which is not only a key measure to balance technological innovation and rights protection, but also an important path to realize the modernization of AI governance.

At present, most academic research on AI regulation focuses on the macro-policy level, with insufficient attention paid to specific operational standards. In the field of data training, existing norms are fragmented^[4]. Industry self-regulatory standards lack enforce-ability, and international standards are difficult to adapt to the local legal environment. This lack of norms leads to high compliance costs for enterprises and insufficient law enforcement basis for regulators, ultimately affecting the healthy

development of AI technology. Therefore, constructing a systematic and enforceable digital standard contract is not only of theoretical innovation value, but also can provide practical solutions for practice. Research on the governance of the artificial intelligence industry stands at a critical transformative juncture. It boasts a wealth of massive application scenarios and clear policy support, yet the current research capacity has not fully translated these advantages into academic and industrial strengths^[5]. Going forward, it is imperative to closely follow the four major trends of scenario-based deep cultivation, standard formulation, paradigm transformation, and ecological collaboration, striving to build a nationally distinctive and influential research hub for applied artificial intelligence governance.

1.2 Construction Principles

The Balance Between Compliance and Development Drawing on the open data-set principles proposed by Mozilla and the requirements of the national AI standard system, the agreement is developed in accordance with five core principles:

1. Legality and Compliance First

Convert requirements such as “informed consent” and “minimum necessary” stipulated in the “Personal Information Protection Law” into mandatory contractual clauses. For instance, personal data used for training must be accompanied by proof of data subject authorization, and sensitive data must complete a Data Protection Impact Assessment (DPIA).

2. Equivalence of Rights and Obligations

Clarify the boundaries of rights and responsibilities among data providers, trainers, and users. This safeguards the data subject’s right to be forgotten”, a built-in data deletion trigger mechanism in the agreement while protecting the algorithmic intellectual property rights of trainers.

3. Technology-Neutral Adaptation

Be compatible with various privacy-enhancing technologies including federated learning and differential privacy. With reference to GB/T 46284-2025 Artificial Intelligence—Technical Specifications for Federated Learning, no specific technical solution is mandated; only security performance indicators are defined.

4. Openness, Collaboration and Sharing

Establish standardized data metadata specifications, support cross-institutional agreement mutual recognition, facilitate the construction of a competitive AI ecosystem, and lower the barrier for small and medium-sized enterprises to participate.

5. Dynamic and Controllable Risks

Introduce a dynamic update mechanism for contractual clauses, enabling real-time optimization of governance rules in response to policy adjustments (e.g., changes in cross-border data transfer rules) and technological developments (e.g., emergence of new attack methods).

2. Concepts and Research Status

2.1 Concepts

As an interdisciplinary subject integrating computer science, mathematics, cognitive science, linguistics and other fields, AI has made remarkable progress in theoretical research and practical application in recent years. Data training, as a key link in the development of AI models, is mainly characterized by the process of data collection, cleaning, annotation and model training, which directly affects the performance and application effect of the model. High-quality data training behavior can significantly improve the performance and generalization ability of the model^[6]. However, data training behavior also faces some challenges, such as data privacy protection and model inter-pretability.

A digital standard contract refers to a contract created, executed, managed and stored using digital technology, containing pre-set clauses, with high consensus, oper-ability and exemplary characteristics, realizing machine readability, executability and automation, and with standardized clauses and data structures. Technically speaking, a digital standard contract integrates part of the mechanism of “technology-driven measures” into the clauses of “legal-driven measures” by using Boolean operators, formulating a “digital standard contract” with high flexibility to meet the needs of multiple scenarios. It can solve complex problems in the process of AI data training, such as fair use, licensing authorization, text and data mining exceptions, rights protection, data sharing, and tort liability division. It can make up for the lack of rules and reconcile the different demands of all stakeholders involved in AI data training to the greatest extent.

AI, data training behavior and digital standard contract have a close synergistic development relationship. The progress of AI technology relies on high-quality data training behavior, and the standardization of data training behavior requires the support of digital standard contracts. Therefore, the formulation and implementation of digital standard contracts can provide a unified specification for the application of AI technology, thus promoting the wide application of AI technology^[7]. On the contrary, the development of AI technology also provides new tools and methods for the formulation of digital standard contracts, such as automatically generating standard clauses through generative AI tools.

The core value of digital standard contract research is reflected in three dimensions: Firstly, it realizes the organic unity of technical ethics and legal compliance by establishing a full-process specification for data collection, processing and use^[8]. Secondly, the pioneering “dynamic compliance” mechanism of the contract enables technical standards to be updated synchronously with legal revisions. In the research of the financial field, clarifying the rights and responsibilities of data providers, algorithm developers and regulators can greatly reduce the blind spots of traditional financial supervision. Finally, the contract innovatively introduces a “technology-legal” dual-track verification system^[9]. Its copyright filtering module not only achieves a higher accuracy of infringement identification, but also forms an evidence fixation standard in line with the judicial interpretation of the Copyright Law.

2.2 Research Status

Current research suffers from the following shortcomings:

1. Disconnection between disciplines.

Scholars in law and ethics lack in-depth understanding of the underlying technical principles of artificial intelligence, Transformer architecture for large models, reinforcement learning^[10]. As a result, their discussions remain superficial and principlebased, fairness should be ensured, failing to propose operable governance solutions that can be embedded into the technical foundation, how to achieve fairness through technical means.

2. Limited practical impact of research outputs. Many findings remain at the level of academic papers and reports, without effective translation into concrete policy provisions, industry standards, or practical guidelines for enterprises.

3. Insufficient industryacademia interaction.

Except for a few leading enterprises such as Yutong, universities and research institutes have limited engagement with most small and mediumsized AI enterprises in the province. Theoretical research has not adequately responded to the practical difficulties encountered by enterprises in their development.

Based on existing research capacity, the regulated development of artificial intelligence industry will exhibit five major trends in the future:

1. From principle discussion to scenariobased deepening: refinement and verticalization. Macrolevel, all-encompassing narratives will be gradually abandoned. Research will deeply dive into specific industry scenarios to produce highly operable governance schemes.

2. From theoretical output to standardsetting: pursuing discourse power and practicality^[11]. Efforts will be actively made to translate theoretical achievements into local standards, industry standards, and even national standards.

3. From humanitiesdriven to technologydriven: paradigm shift in research. More studies will leverage privacy computing, block-chain, federated learning, and other technologies to promote the regulated development of the industry—using technical solutions to address governance challenges. The focus will shift from ex post regulation to ex ante prevention, advancing the implementation of “ethics by design” and “compliance by design” in AI product development workflows.

4. From singleparty dominance to governmentindustryuniversityresearch collaboration: ecological integration and closedloop governance. Industry will provide scenarios and data; enterprises and universities will jointly build laboratories with realworld data and application scenarios for empirical analysis and experimental validation^[12]. Universities will offer theoretical support and solutions while cultivating and delivering interdisciplinary talents for enterprises, forming a closedloop system.

5. From riskfocused to developmentbalanced: shift in research orientation. Research objectives will no longer be limited to mere “control” and “risk avoidance,” but to maximizing the innovative development of the AI industry under effective

governance.

2.3 Research Approach

The construction of digital standard contracts needs to break through the traditional legislative thinking. At the methodological level, an innovative path of “legal encoding” should be adopted to transform abstract legal principles into enforceable technical parameters^[13]. For example, privacy protection requirements can be quantified into specific indicators of data desensitization, and algorithmic fairness can be transformed into threshold standards for model deviation. In addition, in terms of content design, the contract should establish a full-chain normative system: clarify the authorization method and scope restrictions in the data collection stage; stipulate algorithm audit and record retention requirements in the training stage; and establish a liability tracing and relief mechanism in the application stage. This three-dimensional normative framework can effectively cover the entire life cycle of data training.

The design of research methods should focus on the organic combination of empirical and comparative research. By analyzing controversial cases in typical application scenarios such as medical diagnosis and credit evaluation, the legal risk points of data training can be accurately identified. Comparative law research helps to draw on the legislative experience of the EU AI Act, the US Algorithmic Accountability Act and other legislations, and carry out localized adjustment combined with the characteristics of China’s legal system^[14]. The technical verification link needs to design a multi-level test plan, including single-point technical testing, system integration testing and stress testing, to comprehensively evaluate the technical feasibility of the contract. This multi-dimensional research method can ensure the close connection between theoretical construction and practical needs^[15].

The implementation path needs to adopt a progressive strategy. In the short term, industry alliances can be used to promote the formation of best practices, such as establishing an authentication mechanism for the source of training data. In the medium term, a third-party certification system can be developed to grant compliance certification to data training behaviors that meet the standards. In the long run, efforts should be made to promote the upgrading of core standards to mandatory norms and coordinate national legislation through international organizations. In particular, a dynamic update mechanism should be established to enable standard contracts to adapt to the rapid iteration of technology. The standard-setting model of the Internet Engineering Task Force can be used for reference to establish an open and transparent contract revision procedure to ensure that all stakeholders can participate in the improvement of rules^[16].

Where laws remain unclear, “digital standard contracts” can serve directly as an effective tool for resolving complex issues involved in artificial intelligence data training. They meet the urgent need for effective regulation of data training practices by individuals, market regulators, AI R&D and application enterprises (such as Deep Seek), data aggregation platforms, and other stakeholders. The promotion and translation path of digital standard contracts is well-defined and can deliver tangible results. The process of researching and designing the principles, clauses, and other elements of standard contracts also expands the research approaches to core AI-related issues—including subject status, rights and interests, and tort determination—from a brand-new, bottom-up, and highly practice-oriented perspective. This is of great value in promoting the regulated development of the artificial intelligence industry.

3. Problems

The construction of digital standard contracts for AI data training behaviors faces multiple dilemmas. From the perspective of technical characteristics, there is an inherent contradiction between the iterative characteristics of machine learning models and the stability of traditional legal norms. Taking the EU General Data Protection Regulation as an example, its stipulated purpose limitation principle requires that data processing must be consistent with the purpose of initial collection, while machine learning models often derive application scenarios beyond the original scope in the continuous training process^[17]. This tension is particularly prominent in highly sensitive fields such as medical care and finance. Medical institutions often face compliance risks when using patient data for algorithm optimization. In addition, transnational technology companies carrying out data training activities globally need to cope with the differences in data cross-border flow rules among different jurisdictions, which further increases the complexity of legal application.

The lack of algorithmic transparency is a technical bottleneck hindering effective legal regulation^[18,19]. The “black box”

characteristic of deep neural networks makes the model decision-making process difficult to trace. When an algorithm produces discriminatory results, victims often find it difficult to provide evidence. In a 2021 case of algorithmic discrimination on a US recruitment platform, the plaintiff lost the case because they could not obtain the specific parameters of algorithm training. This case highlights the inadequacy of the existing legal relief mechanism^[20]. More notably, with the rise of multi-modal large models, the complexity of training data sources has increased exponentially, making algorithm audits face unprecedented technical challenges. The algorithm impact assessment system under the current legal framework is often difficult to cope with such a large-scale and complex training system.

The identification of data ownership urgently needs legal clarification. Under the traditional copyright law system, the acquisition and use of training data involve multiple rights relationships^[21]. For example, public data crawled by web crawlers may contain copyright-protected content, and the applicable boundary of the current fair use system in machine learning scenarios is not clear. A 2023 infringement lawsuit against a well-known AI painting tool embodies this contradiction, where the plaintiff claimed that their artistic works were included in the training data set without permission^[22]. At the same time, disputes over the ownership of AI-generated content are increasing, and there are obvious differences in judicial practices in various countries. Some countries recognize the copyright of AI-generated content, while others adhere to human author-centered doctrine. This legal uncertainty seriously restricts the rational flow and utilization of data resources.

The existing normative system has structural defects. In terms of normative form, industry self-regulatory standards lack enforceability, while international treaties with legal binding force are difficult to coordinate the interests of various countries. Although the AI ethics guidelines issued by the International Organization for Standardization (ISO) provide a technical framework, they do not involve specific division of legal liability^[23,24]. In terms of normative content, most existing rules focus on the data collection link, and the specification of the training process itself is relatively weak. Taking face recognition technology as an example, although many countries have legislated restrictions on the collection of biometric data, there is a lack of detailed provisions on data processing behaviors in model training. This regulatory gap leads to the lack of clear guidelines for enterprise compliance and also makes regulatory law enforcement face the dilemma of inconsistent standards.

4. Measures

Establish a dynamic update mechanism for contracts It is recommended to set up a permanent revision committee composed of experts in law, computer science, ethics and other disciplines, and formulate a quarterly evaluation cycle to ensure that contract clauses can respond to technological iteration and legal revisions in a timely manner^[25]. At the same time, a rapid transformation mechanism for contract clauses should be established to convert newly promulgated laws and regulations into enforceable technical parameters in the shortest time. This dynamic update mechanism needs to be supported by the development of automated monitoring tools to track the changes of relevant legislation in major jurisdictions around the world in real time.

Standardize transnational data training cooperation At present, there are significant differences in national laws and regulations on data sovereignty, privacy protection and other aspects, which brings compliance challenges to cross-border AI research and development^[26]. It is recommended to develop contract variant modules with regional adaptability on the basis of the existing contract framework. For example, data sovereignty protection clauses can be strengthened for the EU region, and digital copyright protection mechanisms can be focused on for Southeast Asian countries. At the same time, it is necessary to explore the establishment of a transnational certification mechanism and realize the mutual recognition of standards among different jurisdictions through bilateral or multilateral contracts^[27].

Deepen the research on the legal effect of contract execution results. Focus on solving the issue of the admissibility of materials such as algorithm audit reports and compliance certifications in judicial proceedings^[28]. It is recommended that the Supreme People's Court issue relevant judicial interpretations to clarify the criteria for determining the probative force of digital standard contracts in litigation. At the same time, a unified algorithm filing and registration system should be established to provide a basic basis for subsequent liability determination. The improvement of these systems will further enhance the legal authority and enforceability of contracts.

The innovation points of this research mainly lie in the methodological level: First, it proposes a transformation path

of “legal requirements technologization”, decomposing abstract legal principles into quantifiable algorithm parameters, such as concretizing the “minimum necessity principle” in the Personal Information Protection Law into a threshold of access frequency for data fields. Second, it has developed a “compliance embedding” technical architecture, enabling legal requirements to directly influence the neural network training process, achieving a balance between algorithm bias correction and model performance maintenance in tests. Third, it has established an interdisciplinary collaboration mechanism, combining compliance datasets labeled by legal experts with verification tools developed by technical personnel, solving the problem of legal lag behind technological development in traditional governance. These innovations ensure that the protocol maintains the authority of legal norms while also possessing the feasibility of technical solutions.

5. Conclusion

The research points out that digital standard contracts can provide an institutional solution that balances technological innovation and rights protection for AI data training behaviors through the innovative path of “technicalization of legal elements”. Its dynamic update mechanism and cross-border adaptability design are more exemplary for building a global AI governance system.

In the long run, digital standard contracts are expected to develop into a new type of legal infrastructure in the digital era. With the deepening of research, the application scope of digital standard contracts can be expanded from the data training link to the full life cycle management of AI^[29]. In the future, it is possible to explore the integration of the contract framework with emerging technologies such as block-chain and privacy computing to build a more intelligent and transparent compliance governance system. This will not only help regulate the development of AI, but also provide an important reference for the construction of digital rule of law.

Funding

2026 Annual Key Research and Promotion Special Project (Soft Science) of Henan Province {Research on the Digital Standard Contracts for Artificial Intelligence Data Training} 262400410627

Conflict of Interests

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

Reference

- [1] Russell, S. J., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.; J. Yin, Trans.). Tsinghua University Press.
- [2] Guo, D., & Zhang, Y. (2024). Infringement risks of generative AI training data and legal responses. *Journal of Xiangtan University (Philosophy and Social Sciences)*, 48(5), 78–86.
- [3] Jiang, H., & Tan, X. (2026). Beyond techno-nationalism: Governance dilemmas and regulatory approaches to international AI security. *Contemporary China and World*, (1), 65–75.
- [4] Wu, Y., Chen, Y., Yang, H., et al. (2024). Fragmentation of global digital governance under algorithmic institutional competition and its transcendence path. *Modern Distance Education*, 1–15.
- [5] Li, M. (2024). A report on the translation practice of international contract law: AI, data, cybersecurity and the legal landscape from the perspective of teleology [Master’s thesis]. Guangdong University of Foreign Studies.
- [6] Gao, X. (2025). Digital empowerment of dual contract risk control to improve the operational resilience of manufacturing enterprises. *Information Construction*, (6), 56–57.
- [7] He, C., & Ma, G. (2025). Analysis of computer communication technology and electronic information technology in the field of AI. *China New Communications*, 27(20), 16–18.
- [8] Liu, X. (2024). “Non-work use” in generative AI data training and its legitimacy justification. *Legal Forum*, 39(3), 67–78.
- [9] Zuo, S. (2026). Validity identification and dispute resolution path of electronic contracts in the context of digital economy. *Legal Vision*, (3), 52–54.

- [10] Crawford, K. (2021). *The atlas of AI*. Yale University Press.
- [11] Hu, X., & Zhou, Y. (2025). A review of research on economics and management disciplines based on generative AI. *Chinese Journal of Management Science*, 33(1), 76–97.
- [12] Cheng, X. (2026). On the inapplicability of network infringement rules to infringements of generative AI services. *Comparative Law Review*, (1), 125–137.
- [13] Cheng, Y., Chen, G., Chen, H., et al. (2022). Exploration on key technical paths of standard digitization based on AI. *Information Technology and Standardization*, (10), 60–67.
- [14] European Union. (2024). *AI Act (EU AI Act)*.
- [15] China Academy of Information and Communications Technology. (2023). *White paper on AI data training security*.
- [16] Liu, S., Zhou, L., Yang, J., et al. (2022). Evolution of AI industry technology standard cooperation network and subject identification: Based on social network analysis and TOPSIS entropy weight method. *Science and Technology Management Research*, 42(6), 14–152.
- [17] Ma, S., Yi, Z., Pan, G., et al. (2026). Host country data privacy protection policies and China's digital service trade exports: Evidence from the EU General Data Protection Regulation. *Finance and Economics*, 42(1), 39–49.
- [18] European Commission. (2023). *Guidelines on algorithm impact assessment*.
- [19] Gao, Q. (2023). Legal realization path of algorithmic transparency. *Journal of Political Science and Law*, (4), 112–125.
- [20] Wu, Z. (2025). Justification of fair use path for AI data training: A comment on the first fair use cases in China and the United States. *Journal of Shandong University of Science and Technology (Social Sciences)*, 27(6), 53–61.
- [21] Gao, Y. (2024). Regulation of copyright infringement by AI training data. *China Publishing Journal*, (15), 12–18.
- [22] Zhou, H. (2026). Judicial protection of AI models: From the perspective of unfair competition dispute case of “Transformation Comic Special Effect”. *Journal of Law Application*, (3), 71–86.
- [23] Shen, K. (2024). On the implementation mechanism of soft law: Taking AI ethical norms as an example. *Finance and Economics Law Review*, (6), 108–127.
- [24] IEEE. (2021). *Standard for ethical aligned design (IEEE 7000-2021)*.
- [25] Liu, Y. (2026). How to achieve new breakthroughs in China's AI legislation? Practical solutions based on security, rights and governance. *Law and Social Development*, 32(2), 187–207.
- [26] Song, Y. (2025). Legal risks and regulatory paths of generative AI data cross-border flow. *Cybersecurity and Data Governance*, 1–8.
- [27] Cai, X. (2023). A comparative study on AI governance between China, the United States and Europe. *Legal Forum*, (3), 88–95.
- [28] Shen, F. (2025). Risks to judicial justice arising from the development of AI justice and its prevention and governance. *Legal Research*, 1–15.
- [29] Hu, B., Wu, J., & Zhang, S. (2024). Research on AI knowledge creation: Overall framework and future prospects. *Journal of Hangzhou Dianzi University (Social Sciences)*, 20(5), 26–39.

Green Credit Policy and the Sustainability of Green Innovation: Symbolic or Substantive Responses from High-Polluting Firms

Dicheng Wang¹, Qian Zhang², Ziyue Jin³, Han Li^{2*}

1. School of Economics, Gachon University, Seongnam 13120, Republic of Korea

2. School of Business, Gachon University, Seongnam 13120, Republic of Korea

3. School of Graduate, Shenyang Ligong University, Shenyang 110159, China

*Corresponding author: Han Li, powwwer@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: Amid global carbon neutrality and sustainable development goals, and China's transition toward high-quality green and sustainable growth, this study investigates how the Green Credit Policy (GCP) influences the sustainable innovation performance of high-polluting firms. Using Chinese A-share listed firms from 2007 to 2019 and a difference-in-differences (DID) design, we exploit the implementation of the GCP as a quasi-natural experiment to assess its long-term sustainability effects on corporate green transformation. The results reveal that while the GCP significantly promotes symbolic green innovation associated with regulatory compliance, it does not substantially enhance substantive green innovation, raising concerns about the effectiveness of green finance in fostering authentic and high-quality sustainability-oriented innovation. Further analysis shows pronounced heterogeneity across firm types. State-owned enterprises (SOEs) and large firms exhibit improvements in substantive green innovation, thereby contributing more effectively to long-term environmental sustainability and green transformation, whereas non-SOEs and small firms experience tightened financial constraints that crowd out R&D investment, ultimately undermining their sustainable innovation capacity. A series of robustness tests confirms the reliability of these findings. Overall, this study advances the literature on green finance and corporate sustainability by revealing firms' strategic compliance behavior under sustainability-oriented financial regulation, highlighting uneven sustainability outcomes across firm types, and offering policy implications for refining green credit mechanisms to better support genuine green innovation and long-term sustainable development.

Keywords: Green Credit Policy; Sustainable Innovation; Symbolic Green Innovation; Green Technological Transformation; Sustainability; Green Economy

Published: Apr 16, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1266>

1. Introduction

In the context of China's commitment to high-quality green growth and long-term sustainable development, and global agenda for carbon neutrality and sustainability transition, green finance has become a critical policy lever for advancing the economy-wide transition toward environmental sustainability ^[1,2]. Among various green finance instruments, the Green Credit Policy (GCP) represents a cornerstone initiative by embedding environmental sustainability criteria into credit allocation

decisions^[3]. Introduced in 2012 through the Green Credit Guidelines, the policy links firms' financing accessibility to their environmental performance, thereby increasing borrowing costs for high-polluting firms while incentivizing investment in cleaner production and sustainable technological upgrading^[4-6]. By aligning financial resource allocation with ecological priorities, the GCP plays a key role in steering firms toward more sustainable development pathways and long-term environmental outcomes^[7].

Existing research suggests that green credit influences firm behavior through financing constraints, capital costs, and signaling effects^[8-10]. However, most studies emphasize innovation quantity rather than sustainability quality of innovation, overlooking the heterogeneity between substantive and symbolic innovations^[11,12]. In reality, policy-driven innovation may expand in quantity but deteriorate in quality as firms may strategically pursue low-cost, short-term patenting to satisfy compliance requirements. Such behavior risks generating symbolic sustainability outcomes without corresponding improvements in genuine environmental performance, resulting in what has been described as "greenwashing innovation"^[13,14].

Using the GCP as a quasi-natural experiment, this study examines how green credit affects both the quantity and sustainability quality of green innovation among China's heavily polluting firms. We distinguish between substantive innovation and symbolic innovation to assess the policy's true contribution to sustainable technological progress and green transformation. Empirical evidence reveals a clear pattern: while the GCP stimulates more green patenting activity, most of this increase stems from symbolic rather than substantive innovation, suggesting a divergence between formal compliance and substantive sustainability advancement.

This behavior reflects firms' rational responses to regulatory incentives: when policy monitoring focuses on whether firms innovate rather than whether such innovation contributes to long-term sustainability, high-pollution firms tend to favor low-cost, low-risk patenting strategies. Further heterogeneity analysis shows that state-owned and large firms benefit more from green credit access and are more capable of shifting toward substantive green innovation that supports sustainable transformation, whereas non-state and smaller firms face financing barriers and respond primarily through symbolic innovation, thereby experiencing weaker sustainability outcomes.

This study contributes to the literature by (1) differentiating between symbolic and substantive green innovation from a corporate sustainability perspective, (2) providing causal evidence of the GCP's uneven sustainability effects using a DID framework, and (3) offering policy implications to enhance the quality-oriented and sustainability-driven design of China's green finance system.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature and outlines the research questions. Section 3 describes the data, sample selection, and empirical research design. Section 4 presents the main empirical results and robustness tests. Section 5 conducts heterogeneity analyses to explore differential policy effects across firm types. Section 6 concludes the study and discusses the key findings and implications for sustainable development and green finance policy.

2. Theoretical background and research questions

The GCP represents one of the earliest and most influential institutional frameworks within China's green finance system^[15]. As a foundational policy instrument supporting China's transition toward high-quality green growth and long-term sustainable development, the GCP signifies a paradigm shift in financial resource allocation from prioritizing high-growth yet high-emission industries toward encouraging environmentally sustainable, low-carbon, and innovation-driven sectors^[16]. By internalizing environmental considerations into credit allocation decisions, the GCP promotes the alignment of financial development with ecological protection and long-term sustainability outcomes^[17]. Its core objective is to channel financial resources toward low-carbon, energy-efficient, and innovation-driven firms while compelling highly polluting enterprises to transform and upgrade through differentiated credit allocation^[5,18]. In this way, the GCP operationalizes environmental policy through the financial system, using credit as a lever to influence firm behavior^[8,10]. By embedding environmental constraints into credit evaluation and loan approval procedures, the policy turns "green performance" into a critical criterion for obtaining financial support^[19,20], thereby reshaping firms' innovation incentives and investment decisions in ways that support sustainability-oriented technological upgrading^[21,22].

From a theoretical perspective, the GCP draws on both the Porter Hypothesis and the principles of financial constraint theory^[23,24]. According to the Porter Hypothesis, well-designed environmental regulation can stimulate firms to innovate, offsetting compliance costs through efficiency gains and new technologies^[25,26]. The GCP reflects this notion by linking credit access to environmental performance: firms are encouraged to engage in cleaner production, develop green technologies, and upgrade their innovation capabilities to maintain financial competitiveness while advancing the sustainability transition^[27]. At the same time, financial constraint theory suggests that limited access to credit can affect firms' investment behavior, particularly in capital-intensive activities like R&D^[28]. Under the GCP framework, firms with poor environmental compliance experience restricted credit availability and higher borrowing costs, which exert both pressure and motivation to innovate in order to restore creditworthiness and improve sustainability performance^[29]. This dual mechanism of incentive and constraint differentiates green credit from traditional forms of environmental regulation that rely purely on administrative penalties or subsidies^[30].

Existing research indicates that the GCP affects corporate innovation through both positive incentive effects and disciplinary constraint effects^[17,31]. On the one hand, green-oriented firms benefit from lower financing costs, improved risk ratings, and increased access to financial markets, enabling them to expand their R&D investment and pursue long-term innovation projects with greater sustainability value^[18]. Empirical evidence shows that credit incentives under green finance frameworks can effectively promote R&D spending and enhance firms' capacity for technological upgrading^[30]. On the other hand, highly polluting firms face stricter credit supervision and financing pressures, which may initially constrain their operations but also motivate them to improve environmental performance and technological capability to regain credit access^[9,21]. Through this process, the GCP creates a self-reinforcing mechanism in which financial constraints encourage green transformation by altering the cost–benefit structure of pollution and innovation. Therefore, it is reasonable to expect that the GCP contributes positively to the green innovation activities of highly polluting enterprises and their sustainability-oriented upgrading. Accordingly, we propose the following research question:

RQ1: Does the Green Credit Policy significantly enhance the green innovation of highly polluting firms?

While the implementation of the GCP is expected to stimulate overall green innovation, an increase in the quantity of innovations does not necessarily translate into an improvement in their quality or their sustainability relevance^[8,22]. Quantity-based measures such as patent counts may overestimate the effectiveness of environmental policies if the underlying innovations lack technological substance or market applicability and fail to generate measurable sustainability outcomes^[20]. According to the Porter Hypothesis, only well-designed and flexible environmental regulations can induce high-quality innovation through the so-called innovation compensation effect, in which firms compensate for regulatory costs by developing new technologies that enhance efficiency and competitiveness while supporting the sustainability transition^[26]. However, in practice, financing constraints, short-term performance assessments, and regulatory compliance pressures may distort firms' innovation behavior, driving them to pursue compliance-oriented innovation rather than genuine technological breakthroughs that deliver long-term environmental and sustainability benefits^[32].

For heavily polluting firms, the GCP functions as both a constraint and a signal^[33]. On one side, it raises financing costs and limits long-term credit availability, making sustained investment in risky, high-value innovation projects more difficult^[34]. On the other side, it signals governmental and societal expectations for firms to demonstrate “green behavior” through visible outputs such as green patents^[35]. Under this dual pressure, firms may strategically shift their innovation strategy toward low-cost, low-risk patenting activities that satisfy external evaluation metrics but lack substantive technological value and contribute little to long-term sustainable development. This leads to what recent studies have described as a quantity–quality mismatch in green innovation outcomes, where the number of green patents rises while their novelty and impact decline while limited improvements in sustainability performance^[36].

Building on this logic, this study extends the existing literature by investigating the heterogeneous effects of the GCP on different types of green innovation and their implications for sustainable development^[37]. Specifically, we distinguish between substantive green innovation and symbolic green innovation. Substantive green innovation refers to technological activities that lead to genuine environmental and technological progress, typically represented by green invention patents^[33]. These

patents involve higher R&D input, longer development cycles, and greater uncertainty, but they contribute meaningfully to firms' long-term competitiveness and environmental performance and thereby strengthen sustainability outcomes^[30]. Symbolic green innovation, in contrast, refers to innovations that are primarily quantitative and strategic, represented by green utility model patents. These innovations often require less investment and are pursued to demonstrate compliance with environmental regulations or to improve firms' public image without generating substantial sustainability impact or yielding substantial technological breakthroughs^[29,38].

This differentiation is important because it captures how firms internalize environmental financial policies in their innovation behavior and whether green finance drives substantive sustainability transformation. If the GCP primarily stimulates symbolic innovation, it may indicate that the policy is effective in promoting visible but superficial compliance; however, if it encourages substantive innovation, it would confirm that green finance has succeeded in driving meaningful technological transformation consistent with the sustainability transition. The distinction between these two innovation types also aligns with the broader debate in sustainability research regarding greenwashing behaviors—where firms adopt the appearance of environmental responsibility without significant environmental impact improvement or sustained progress toward sustainable development goals. Accordingly, this study proposes the following research questions to guide the empirical analysis:

RQ2: Does the Green Credit Policy promote substantive green innovation among firms?

RQ3: Does the Green Credit Policy promote symbolic green innovation among firms?

3. Sample and research design

3.1 Sample and Data

This study employs a panel dataset of Chinese A-share listed firms spanning the period from 2007 to 2019. Owing to the disruptions introduced by the COVID-19 pandemic, the most recent observations were excluded from the sample^[39]. Data on firms' green patents are obtained from the China Research Data Services (CNRDS) platform, while financial and accounting indicators are primarily drawn from the China Stock Market & Accounting Research (CSMAR) database. Following established practices in prior studies, we exclude several types of firms to ensure sample consistency and data reliability^[40]. Specifically, we remove: (1) listed companies engaged in monetary and financial services, such as banking and insurance; (2) ST, ST, and PT firms that are subject to special treatment due to abnormal financial conditions; (3) firms with a leverage ratio greater than 1 or less than 0; and (4) firms with severe data deficiencies or missing key variables.

After applying these screening criteria, the final sample consists of 10,489 firm-year observations. To mitigate the influence of outliers and enhance the robustness of the results, all continuous variables are winsorized at the 1st and 99th percentiles^[41]. This process minimizes the impact of extreme values and ensures the reliability and stability of subsequent empirical analyses.

3.2 Research Design

3.2.1 Dependent Variable

The dependent variable in this study is firms' green innovation, which is further divided into three dimensions to capture heterogeneity in innovation characteristics: substantive green innovation, symbolic green innovation, and overall green innovation level. (1) Substantive green innovation is measured by the number of green invention patent applications filed by a firm in a given year (G_inva). (2) Symbolic green innovation is measured by the number of green utility model patent applications filed by a firm in the same period (G_uma). (3) Green innovation level represents the total number of green patent applications, including both invention and utility model patents (G_total).

To address potential issues of heteroscedasticity and skewed distribution in patent data, the quantities of all three types of patents are transformed by taking the natural logarithm of $(1 + \text{number of patents})$ ^[42]. This logarithmic adjustment helps stabilize variance and allows for more consistent statistical estimation across firms and years.

3.2.2 Independent Variable

The core independent variable of this study is the interaction term between the policy shock and the treatment group—denoted as $(\text{Treat}_i \times \text{Post}_i)$ —which captures the net impact of the GCP on the green innovation of high-polluting firms.

The treatment group dummy variable (Treat_i) is defined according to the Guidelines for the Classification of Listed

Companies' Industries. Firms operating in 19 high-polluting industries—including thermal power, iron and steel, and chemical engineering—are assigned a value of 1, while firms in non-polluting industries are assigned 0.

The policy timing dummy variable ($Post_t$) equals 1 for the years 2012 and beyond, representing the period after the introduction of the Green Credit Guidelines, and 0 for years prior to 2012.

The interaction term ($Treat_t \times Post_t$) takes the value of 1 when a firm belongs to a high-polluting industry ($Treat_t = 1$) during or after 2012 ($Post_t = 1$), and 0 otherwise. The coefficient of this interaction term reflects the causal effect of the GCP on the green innovation activities of high-polluting firms, isolating the policy's impact from temporal and sectoral heterogeneity.

3.2.3 Control Variables

To mitigate potential endogeneity and omitted-variable bias, this study includes a comprehensive set of firm-level control variables that may influence green innovation performance^[43].

We control for Tobin's Q (Tobin Q) to capture firm market valuation and Return on Assets (ROA) to measure profitability. Leverage ratio (Lev) reflects the firm's capital structure, while the book-to-market ratio (BM) represents valuation relative to underlying fundamentals. Furthermore, firm size (Size) and firm age (Age) are included to account for heterogeneity in firm scale and maturity, both of which may affect firms' innovation capacity and risk tolerance. Property, Plant, and Equipment ratio (PPE) captures asset intensity, R&D intensity (R&D) measures the proportion of resources devoted to innovation activities, and cash ratio (Cash) reflects liquidity conditions.

All control variables are defined consistently with prior empirical studies on green finance and corporate innovation. Detailed definitions and measurement descriptions of all variables are presented in Table 1.

Table 1. Variable definitions.

Variable	Definition
G_inva	green invention patents
G_uma	green utility model patents
G_total	total green patents
treat	1 for high-polluting firms, 0 for non- high-polluting firms
post	1 if years from 2012 onward, otherwise 0
Treat*Post	Interaction term of the DID treatment dummy variable
Tobin Q	Market value of equity plus total liabilities divided by total assets
Roa	Net profit divided by total assets
Lev	Total liabilities divided by total assets
BM	Book-to-market ratio
Size	Natural logarithm of total assets
Age	Natural logarithm of years since firm establishment
PPE	Net value of fixed assets divided by total assets
R&D	R&D expenditure divided by operating revenue
Cash	Cash holdings divided by total assets

3.2.4 Descriptive statistics

Table 2 reports the descriptive statistics for 10,489 observations. The mean values of green invention patents (G_inva), green utility model patents (G_uma), and total green patents (G_total) are 0.790, 0.765, and 1.108, respectively, with large dispersions, suggesting low but heterogeneous green innovation performance. Among policy variables, 37.9% of firms are in the treatment group, 61.6% fall within the post-policy period, and the interaction term has a mean of 0.233, ensuring adequate variation for policy identification. Regarding firm characteristics, the mean Tobin's Q is 1.933, ROA is 0.043, and leverage is 0.484, indicating moderate profitability and leverage. Average firm size is 22.48, while R&D intensity (0.013) remains low, implying limited innovation investment among high-polluting firms. Overall, the results indicate that green innovation levels are modest but heterogeneous, and the GCP provides a suitable setting to examine differential innovation responses.

Table 2. Descriptive statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
G_inva	10489	0.790	1.239	0	5.056
G_uma	10489	0.765	1.219	0	4.736
G_total	10489	1.108	1.477	0	5.587
Treat	10489	0.379	0.485	0	1
Post	10489	0.616	0.486	0	1
Treat*Post	10489	0.233	0.423	0	1
Tobin Q	10489	1.933	1.196	0.855	7.722
Roa	10489	0.043	0.047	-0.102	0.204
Lev	10489	0.484	0.188	0.075	0.860
BM	10489	0.313	0.143	0.071	0.758
Size	10489	22.48	1.349	19.90	26.39
Age	10489	2.775	0.351	1.792	3.401
PPE	10489	0.255	0.182	0.002	0.752
R&D	10489	0.013	0.023	0	0.119
Cash	10489	0.589	0.831	0.029	5.565

3.3 Research Design

In line with the research purpose of this paper and the aforementioned theoretical analysis, we construct the following baseline model using the difference-in-differences method:

$$greenpatent_{it} = \beta_0 + \beta_1 treat_i + \beta_2 post_t + \beta_3 treat_i * post_t + \gamma Control_{it} + Firm_i + Year_t + \varepsilon_{it} \quad (1)$$

4. Empirical results

4.1 Baseline results

Table 3 reports the benchmark regression results examining the impact of the GCP on the green innovation activities of high-polluting firms. The policy effect is captured by the did interaction term, with both firm and year fixed effects included to control for unobserved heterogeneity and time-specific influences. Columns (1) – (3) present results without control variables, while Columns (4) – (6) include firm-level controls. The dependent variables are G_inva, G_uma, and G_total, respectively.

In the baseline model without controls, the coefficient of did is G_inva (−0.0352, $p > 0.1$), for G_uma (0.1065, $p < 0.05$), and G_total (0.0830, $p > 0.1$). The positive and significant coefficient for G_uma at the 5% level indicates that the GCP significantly promotes the application of green utility model patents. In contrast, the coefficients for G_inva and G_total are statistically insignificant, suggesting that the policy has little influence on either substantive innovation or total green innovation output.

After including control variables, the results remain consistent. The coefficient for G_uma (0.1184, $p < 0.05$) increases slightly, confirming the policy's positive impact on incremental green innovation. The coefficient for G_inva (−0.0163, $p > 0.1$) remains negative and statistically insignificant, implying that the GCP does not significantly enhance high-quality innovation. The coefficient for G_total (0.1035, $p < 0.1$), this improvement is mainly driven by the expansion of G_uma, indicating that the increase in total green patents stems primarily from low-cost, short-cycle innovations.

These results demonstrate that the GCP primarily stimulates symbolic or incremental green innovation, rather than substantive technological innovation. The findings support the view that high-polluting firms, under financial constraints and regulatory pressure, prefer low-cost, compliance-oriented innovation instead of engaging in high-risk, technology-intensive R&D.

Table 3. Benchmark regression results.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	G_inva	G_uma	G_total	G_inva	G_uma	G_total
treat*post	-0.0352 (0.0544)	0.1065** (0.0521)	0.0830 (0.0613)	-0.0163 (0.0528)	0.1184** (0.0510)	0.1035* (0.0591)
Tobin Q				0.0062 (0.0175)	-0.0137 (0.0166)	-0.0153 (0.0191)
Roa				-0.2963 (0.3116)	-0.1412 (0.3135)	-0.3002 (0.3705)
Lev				-0.1979 (0.1831)	-0.1284 (0.1786)	-0.2307 (0.2066)
BM				0.0581 (0.1857)	-0.0715 (0.1879)	-0.1125 (0.2104)
Size				0.3642*** (0.0468)	0.2884*** (0.0451)	0.4107*** (0.0526)
Age				0.3710 (0.2274)	0.1494 (0.2247)	0.3868 (0.2471)
PPE				0.1201 (0.1559)	0.4369*** (0.1595)	0.2670 (0.1871)
R&D				5.5574*** (1.0434)	3.0497*** (1.0230)	5.4147*** (1.0879)
Cash				-0.0540*** (0.0175)	-0.0148 (0.0178)	-0.0456** (0.0203)
Constant	0.2360*** (0.0260)	0.2079*** (0.0257)	0.3411*** (0.0304)	-8.4758*** (1.0426)	-6.4055*** (0.9986)	-9.3541*** (1.1335)
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	10,489	10,489	10,489	10,489	10,489	10,489
R-squared	0.2125	0.2029	0.2662	0.2507	0.2250	0.3000

Note: ***, **, and * indicate 10%, 5%, and 1% significance levels, respectively. Robust standard errors in parenthesis.

4.2 Robustness checks

To further verify the reliability of the benchmark regression results and the validity of causal inference, this section systematically rules out potential confounding factors through multiple robustness tests, including parallel trend test, Placebo test, counterfactual tests, Propensity Score Matching-Difference-in-Differences (PSM-DID), replacement of dependent variables, and lagged independent variables.

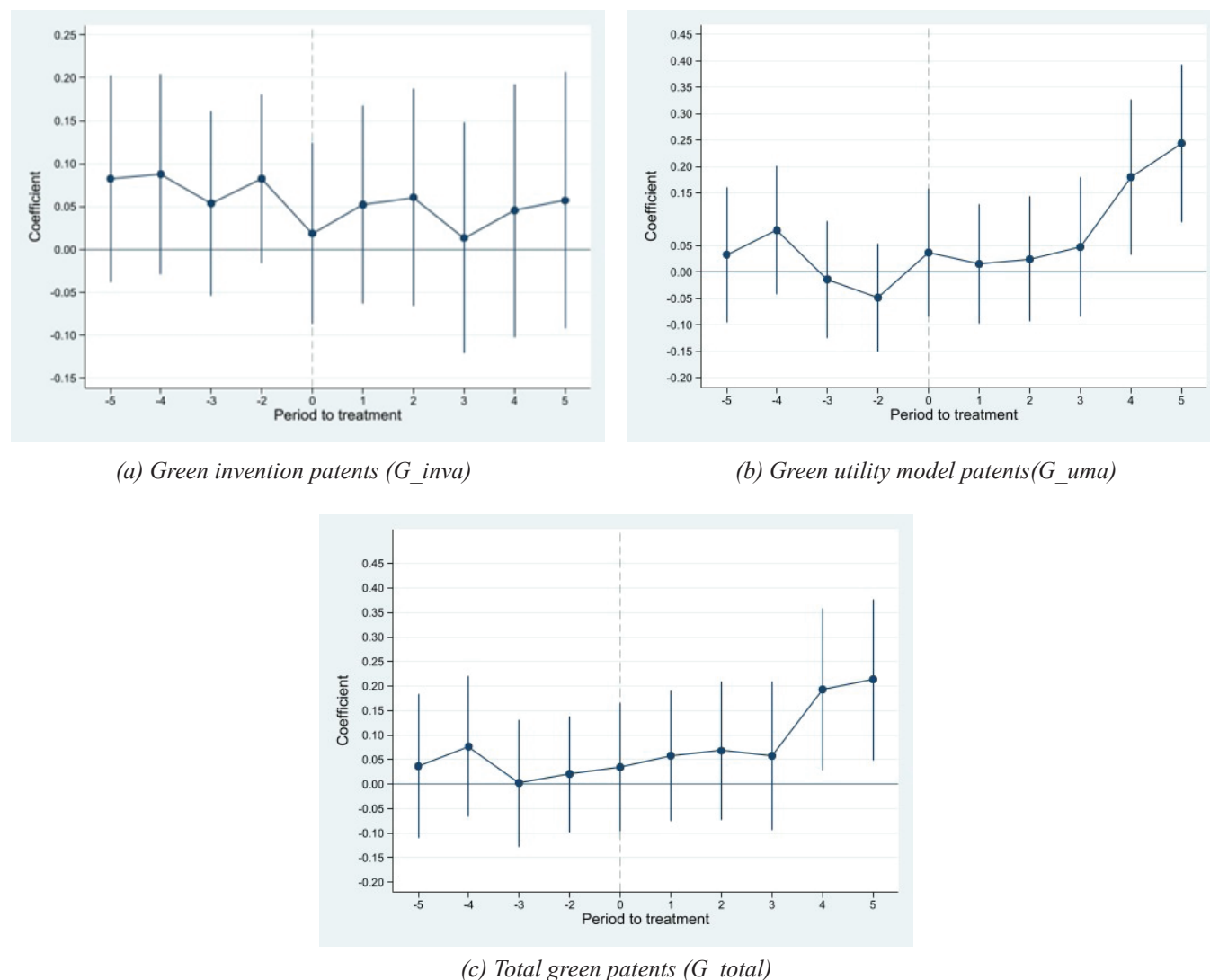
4.2.1 Parallel trend test

Passing the parallel trend test is one of the important conditions for satisfying the $Treat_i \times Post_t$ premise. According to the methods of Liu and Qiu ^[44], this paper introduces the continuous $Treat_i \times Post_t$ model for parallel trend test. The specific model settings are as follows:

$$greenpatent_{it} = \alpha_0 + \sum_{t=2019, t \neq 2011}^{2007} \alpha_t treat_i * post_t + \gamma Control_{it} + Firm_i + Year_t + \varepsilon_{it} \quad (2)$$

Fig. 1 presents the parallel trend test results for the three types of green innovation indicators: G_inva , G_uma , and G_total . Before the implementation of the GCP, the estimated coefficients for all three indicators fluctuate narrowly around zero without a significant upward or downward trend, indicating that the treatment and control groups followed a similar trajectory prior to the policy. This finding satisfies the parallel trend assumption required for the DID framework, supporting the credibility of subsequent causal inference. After the policy implementation, the coefficients of G_uma and G_total show a mild but steady increase, whereas G_inva remains largely unchanged, suggesting that the observed post-policy rise in total green patents is primarily driven by growth in green utility model patents rather than substantive invention patents. This pattern preliminarily implies that the GCP may have encouraged symbolic rather than substantive green innovation among heavily polluting firms.

Figure 1. Parallel trend test.



4.2.2 Placebo test

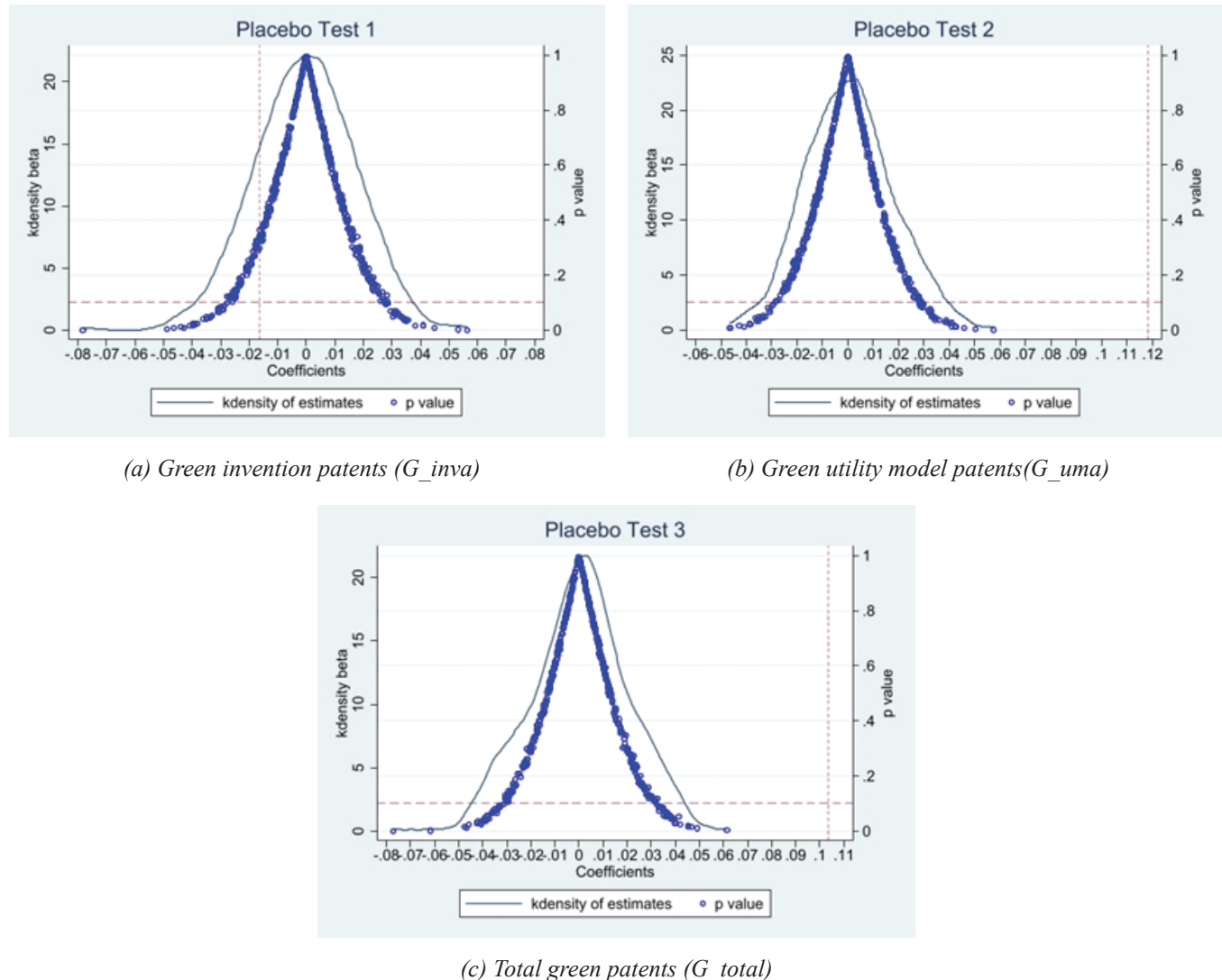
To verify that the impact of the GCP on corporate green innovation is not driven by random factors, we conduct a placebo test to mitigate potential bias from omitted variables. Specifically, 500 random samples are drawn from the full dataset to generate pseudo-policy dummy variables ($Treat_c \times Post_t$), and the baseline model (1) is re-estimated accordingly. Since the placebo treatment is randomly assigned, its estimated coefficients should not systematically deviate from zero.

As shown in Fig. 2, the results support this expectation. For G_uma and G_total , the average estimated coefficients of the placebo variables are close to zero and far smaller than the corresponding baseline estimates. The distributions are approximately normal, with most p-values exceeding 0.10, indicating a lack of statistical significance. In contrast, the benchmark coefficient of G_inva lies within the high-density region of the placebo estimates, suggesting that the observed

effect of the GCP on substantive innovation may be partly influenced by random variation.

These results reinforce the robustness of the baseline findings. The GCP has a statistically reliable effect on green utility model patents and total green patents, whereas its impact on green invention patents is less stable, consistent with the interpretation that the policy primarily stimulates symbolic rather than substantive green innovation among heavily polluting firms.

Figure 2. Placebo test.



4.2.3 Counterfactual test

To eliminate concerns about spurious time shocks, we construct a counterfactual interaction term ($P.treat*post$) and regress it on the green innovation indicators. As reported in Table 4 (columns 1–3), the coefficients of $P.treat*post$ for G_inva (-0.0293), G_uma (0.0633), and G_total (0.0673) are all statistically insignificant ($p > 0.1$). These results indicate no differential pre-policy trends in green innovation between treatment and control groups when the actual 2012 policy is excluded, confirming that the baseline effects are indeed driven by the implementation of the GCP rather than by random or time-specific factors.

4.2.4 PSM-DID

To address potential selection bias, we employ nearest-neighbor matching to construct a balanced treatment and control sample. As reported in Table 4 (columns 4–6), the DID coefficient remains significant for G_uma (0.1144, $p < 0.05$) and G_total (0.1055, $p < 0.1$), but insignificant for G_inva (-0.0135, $p > 0.1$). The matched samples also yield higher R^2 values (0.680–0.729), indicating improved model fit. These results confirm that the finding the GCP promote symbolic rather than substantive green innovation is robust after accounting for differences in firm characteristics between groups.

Table 4. Counterfactual Test and PSM-DID.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	G_inva	G_uma	G_total	G_inva	G_uma	G_total
P. treat*post	-0.0293 (0.0528)	0.0633 (0.0530)	0.0673 (0.0613)			
treat*post				-0.0135 (0.0529)	0.1144** (0.0513)	0.1055* (0.0593)
Tobin Q	0.0060 (0.0175)	-0.0143 (0.0166)	-0.0156 (0.0190)	0.0034 (0.0176)	-0.0121 (0.0171)	-0.0164 (0.0193)
Roa	-0.2972 (0.3120)	-0.1201 (0.3139)	-0.2829 (0.3707)	-0.3180 (0.3210)	-0.2156 (0.3247)	-0.3625 (0.3830)
Lev	-0.1999 (0.1829)	-0.1207 (0.1788)	-0.2234 (0.2063)	-0.1915 (0.1841)	-0.1062 (0.1838)	-0.2000 (0.2092)
BM	0.0568 (0.1856)	-0.0560 (0.1871)	-0.0994 (0.2094)	0.0239 (0.1859)	-0.1079 (0.1910)	-0.1476 (0.2106)
Size	0.3641*** (0.0468)	0.2853*** (0.0450)	0.4084*** (0.0525)	0.3667*** (0.0475)	0.2903*** (0.0457)	0.4111*** (0.0533)
Age	0.3732 (0.2274)	0.1628 (0.2248)	0.3962 (0.2471)	0.2411 (0.2261)	0.0921 (0.2272)	0.2466 (0.2463)
PPE	0.1200 (0.1564)	0.4476*** (0.1603)	0.2756 (0.1880)	0.1268 (0.1568)	0.4382*** (0.1627)	0.2676 (0.1887)
R&D	5.5564*** (1.0430)	3.0368*** (1.0216)	5.4051*** (1.0854)	5.8030*** (1.0639)	3.4988*** (1.0360)	5.7603*** (1.1092)
Cash	-0.0539*** (0.0175)	-0.0145 (0.0179)	-0.0455** (0.0204)	-0.0554*** (0.0181)	-0.0129 (0.0180)	-0.0443** (0.0210)
Constant	-8.4767*** (1.0425)	-6.3819*** (0.9966)	-9.3349*** (1.1308)	-8.0941*** (1.1164)	-6.0653*** (1.0670)	-8.7573*** (1.2086)
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	10,489	10,489	10,489	10,178	10,178	10,178
R-squared	0.2507	0.2241	0.2995	0.7055	0.6807	0.7291

Note: ***, **, and * indicate 10%, 5%, and 1% significance levels, respectively. Robust standard errors in parenthesis.

4.2.5 Dependent variable replacement

The baseline analysis employs green patent applications, which may overstate actual innovation activity. To address potential measurement bias, we replace the dependent variables with green patent grants (G_inva_grant, G_uma_grant, G_total_grant) to capture realized innovation output. As shown in Table 5 (columns 1–3), the DID coefficient remains significant for G_uma_grant (0.0918, $p < 0.1$) and G_total_grant (0.1166, $p < 0.05$), but insignificant for G_inva_grant (-0.0265, $p > 0.1$). These results eliminate concerns of inflated patent applications and confirm the robustness of the baseline findings demonstrating that the GCP primarily enhances symbolic rather than substantive green innovation.

Enterprise innovation is influenced by multiple factors, and if relevant variables are omitted from the model, potential

endogeneity bias may arise. This study examines the extent of omitted variable bias by calculating the ratio of coefficient differences between models with and without observable controls. This ratio serves as an indicator of the magnitude of unobservable bias relative to observable factors.

Table 5. Dependent Variable Replacement.

VARIABLES	(1)	(2)	(3)
	G_inva_grant	G_uma_grant	G_total_grant
treat*post	-0.0265 (0.0413)	0.0918* (0.0488)	0.1166** (0.0546)
Tobin Q	0.0041 (0.0132)	-0.0068 (0.0161)	-0.0059 (0.0182)
Roa	-0.5944** (0.2377)	-0.3667 (0.2963)	-0.6845** (0.3162)
Lev	-0.1010 (0.1348)	-0.0837 (0.1689)	-0.2033 (0.1889)
BM	0.1118 (0.1469)	-0.0267 (0.1781)	-0.0568 (0.1915)
Size	0.1834*** (0.0324)	0.2695*** (0.0430)	0.3173*** (0.0461)
Age	0.3472* (0.1959)	0.2503 (0.2165)	0.3881* (0.2344)
PPE	0.2693** (0.1157)	0.4628*** (0.1495)	0.4436*** (0.1659)
R&D	4.3834*** (0.7954)	3.4581*** (0.9897)	5.5630*** (0.9945)
Cash	-0.0312** (0.0148)	-0.0075 (0.0167)	-0.0318 (0.0193)
Constant	-4.7534*** (0.7839)	-6.2954*** (0.9574)	-7.5259*** (1.0215)
Firm FE	YES	YES	YES
Year FE	YES	YES	YES
Observations	10,489	10,489	10,489
R-squared	0.1594	0.2016	0.2542

Note: ***, **, and * indicate 10%, 5%, and 1% significance levels, respectively. Robust standard errors in parenthesis.

5. Heterogeneity analysis

5.1 Firm ownership

Table 6 presents the heterogeneity analysis by ownership structure, dividing the sample into SOEs and non-SOEs. Columns (1)–(3) report the results for SOEs, while columns (4)–(6) correspond to non-SOEs. For SOEs, the DID coefficients are significantly positive across all three innovation indicators: G_inva (0.1117, $p < 0.1$), G_uma (0.2304, $p < 0.01$), and G_total (0.2588, $p < 0.01$). These findings indicate that the GCP has a clear and positive influence on both substantive and symbolic green innovation among SOEs ^[45]. The stronger responses of SOEs likely reflect their closer alignment with government

environmental objectives, higher compliance incentives, and preferential access to bank financing^[46]. In addition, SOEs generally possess more abundant R&D resources and stronger technological capabilities, which enable them to transform financial incentives into tangible innovation outcomes^[30]. Therefore, the GCP strengthens both policy compliance motivation and innovation-driven competitiveness among SOEs.

In contrast, the results for non-SOEs show an opposite pattern. The DID coefficient for *G_inva* (-0.2290, $p < 0.01$) is significantly negative, and the coefficient for *G_total* (-0.1598, $p < 0.05$) is also negative, while that for *G_uma* (-0.0753) is statistically insignificant. These results suggest that non-SOEs experience a decline in substantive green innovation and obtain little benefit from the GCP. Because non-SOEs rely more on market-based financing and lack the implicit guarantees enjoyed by SOEs, they face tighter credit constraints under the GCP framework^[9]. Consequently, these firms tend to reduce investment in costly long-term R&D projects, particularly in green invention patents, and instead prioritize short-term compliance and operational stability. Overall, these findings highlight the asymmetric impact of the GCP across ownership types, revealing that SOEs leverage policy support to enhance innovation, whereas non-SOEs are more vulnerable to financial crowding-out effects.

Table 6. SOEs and non-SOEs.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	G_inva	G_uma	G_total	G_inva	G_uma	G_total
treat*post	0.1117* (0.0665)	0.2304*** (0.0642)	0.2588*** (0.0723)	-0.2290*** (0.0831)	-0.0753 (0.0791)	-0.1598* (0.0966)
Tobin Q	0.0388* (0.0235)	-0.0094 (0.0211)	0.0106 (0.0249)	-0.0360 (0.0257)	-0.0262 (0.0257)	-0.0516* (0.0291)
Roa	-0.3894 (0.4073)	-0.4488 (0.3854)	-0.4037 (0.4474)	-0.3511 (0.4723)	0.0723 (0.4912)	-0.3521 (0.5942)
Lev	-0.0212 (0.2428)	-0.2571 (0.2371)	-0.1001 (0.2639)	-0.6276** (0.2786)	-0.2191 (0.2752)	-0.6526** (0.3303)
BM	0.3367 (0.2401)	-0.0666 (0.2433)	0.1073 (0.2645)	-0.4859* (0.2869)	-0.2772 (0.2849)	-0.6086* (0.3464)
Size	0.3566*** (0.0609)	0.3038*** (0.0581)	0.4076*** (0.0671)	0.4193*** (0.0734)	0.3164*** (0.0704)	0.4667*** (0.0834)
Age	0.4706 (0.3097)	0.4802 (0.3012)	0.6030* (0.3249)	0.2480 (0.3406)	-0.2200 (0.3331)	0.1312 (0.3818)
PPE	-0.0347 (0.2048)	0.3702* (0.2024)	0.1318 (0.2370)	0.2887 (0.2292)	0.4358* (0.2548)	0.3728 (0.3023)
R&D	6.4377*** (1.5816)	2.8666** (1.4547)	5.9429*** (1.5956)	5.3924*** (1.3760)	3.7605*** (1.4041)	5.6374*** (1.4750)
Cash	-0.0493** (0.0250)	-0.0220 (0.0229)	-0.0410 (0.0263)	-0.0672*** (0.0229)	-0.0235 (0.0256)	-0.0637** (0.0290)
Constant	-8.8059*** (1.4244)	-7.5002*** (1.3409)	-10.0202*** (1.4738)	-8.8161*** (1.5217)	-5.9174*** (1.4302)	-9.4056*** (1.7133)
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	6,589	6,589	6,589	3,900	3,900	3,900
R-squared	0.2637	0.2434	0.3208	0.2475	0.2099	0.2836

Note: ***, **, and * indicate 10%, 5%, and 1% significance levels, respectively. Robust standard errors in parenthesis.

5.2 Firm size

Table 7 reports the heterogeneity analysis based on firm size, dividing the sample in-to large firms and small firms. Columns (1)–(3) show the results for large firms, while columns (4)–(6) correspond to small firms. For large firms, the DID coefficients are significantly positive for G_uma (0.2595, $p < 0.01$) and G_total (0.2152, $p < 0.05$), but insignificant for G_inva (0.0114, $p > 0.1$). These results indicate that the GCP stimulates the overall green innovation of large firms mainly through symbolic rather than substantive innovation.

Large enterprises generally possess stronger R&D capacity, more stable financial structures, and greater adaptability to policy changes^[30]. With easier access to external financing and closer relationships with financial institutions, they are better able to respond to regulatory requirements^[21]. To demonstrate environmental commitment and maintain policy compliance, large firms tend to increase low-cost utility model patents rather than engage in high-risk invention projects^[29]. Their green innovation activities therefore expand in scale but not in technological depth, suggesting that the GCP promotes compliance-oriented behavior rather than genuine technological advancement among large firms.

In contrast, small firms exhibit negative and weaker effects. The DID coefficients for G_inva (-0.1219, $p < 0.05$), G_uma (-0.1015, $p < 0.1$), and G_total (-0.0923, $p > 0.1$) suggest that the GCP constrains rather than encourages their innovation. Due to resource limitations and tighter financing conditions^[47], small firms often cut R&D spending to pre-serve liquidity. These findings imply that the GCP imposes heavier financial and adjustment pressures on smaller enterprises, reducing their incentives to pursue both substantive and symbolic green innovation^[9].

Table 7. Firm size.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	G_inva	G_uma	G_total	G_inva	G_uma	G_total
treat*post	0.0114 (0.0801)	0.2595*** (0.0796)	0.2152** (0.0876)	-0.1219** (0.0617)	-0.1015* (0.0540)	-0.0923 (0.0707)
Tobin Q	0.0631 (0.0383)	0.0373 (0.0346)	0.0373 (0.0406)	-0.0335* (0.0186)	-0.0539*** (0.0175)	-0.0590*** (0.0204)
Roa	-0.9296* (0.5273)	-0.6742 (0.5256)	-0.8480 (0.6000)	-0.0733 (0.3580)	0.1035 (0.3546)	-0.1125 (0.4389)
Lev	-0.6762** (0.3414)	-0.4235 (0.3205)	-0.6106* (0.3598)	-0.2490 (0.2058)	-0.2784 (0.2039)	-0.3865 (0.2429)
BM	0.2183 (0.3236)	0.1334 (0.3046)	0.1382 (0.3398)	-0.5342*** (0.1936)	-0.6459*** (0.1966)	-0.8230*** (0.2294)
Size	0.3740*** (0.0727)	0.2896*** (0.0734)	0.4084*** (0.0826)	0.3252*** (0.0554)	0.2604*** (0.0498)	0.3824*** (0.0610)
Age	0.4149 (0.3494)	0.1435 (0.3507)	0.3514 (0.3716)	0.6277** (0.2611)	0.4169* (0.2500)	0.7162** (0.3013)
PPE	0.0756 (0.2526)	0.2997 (0.2521)	0.1795 (0.2957)	-0.0962 (0.1603)	0.3088* (0.1687)	0.0643 (0.2020)
R&D	6.5302*** (2.1119)	3.5593* (2.0764)	5.4901** (2.1905)	5.7258*** (1.1053)	3.3159*** (1.1078)	6.0229*** (1.1607)
Cash	-0.0713* (0.0397)	-0.0363 (0.0430)	-0.0501 (0.0469)	-0.0466*** (0.0173)	-0.0071 (0.0177)	-0.0406* (0.0207)
Constant	-8.8684*** (1.7567)	-6.5142*** (1.7747)	-9.3623*** (1.9453)	-7.7265*** (1.1070)	-5.9432*** (1.0309)	-8.8973*** (1.2043)
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Observations	5,172	5,172	5,172	5,317	5,317	5,317
R-squared	0.3214	0.3009	0.3751	0.1986	0.1616	0.2394

Note: ***, **, and * indicate 10%, 5%, and 1% significance levels, respectively. Robust standard errors in parenthesis.

6. Conclusion and Recommendation

6.1 Conclusions

This study examines the impact of China's GCP on the green innovation behavior of heavily polluting firms and its implications for the sustainability transition under carbon neutrality goals. By distinguishing between substantive and symbolic green innovation, the analysis offers a more nuanced understanding of how sustainability-oriented financial regulation shapes not only the scale but also the quality of corporate innovation that underpins the sustainability transition. Employing a difference-in-differences approach and conducting heterogeneity analyses across firm ownership and firm size, this study identifies the causal effects of the GCP and reveals significant variation in firms' innovation responses and sustainability-relevant outcomes.

The empirical results indicate that the GCP significantly increases total green patent output and primarily promotes symbolic green innovation, while its effect on substantive green innovation remains statistically insignificant. This suggests that, although the policy is effective in expanding the scale of green innovation activities, it falls short in fostering deeper, technology-intensive innovation that contributes to genuine and long-term sustainability outcomes, including meaningful emissions reduction and cleaner production improvements. Further heterogeneity analyses show that SOEs respond most strongly to the GCP by improving both substantive and symbolic green innovation, likely due to their preferential access to credit resources and stronger incentives to comply with policy objectives and demonstrate sustainability performance. Large firms also exhibit a positive response, but this response is mainly symbolic, reflecting a tendency to engage in low-cost patenting strategies to meet regulatory expectations without proportional improvements in sustainability impact. In contrast, non-SOEs and small firms experience tighter financing constraints and limited innovation capacity, which discourages substantive green innovation and may even crowd out high-quality technological investment, thereby weakening their sustainable development capacity.

These findings highlight a potential unintended consequence of green credit policies: heavily polluting firms may prioritize short-term, low-cost compliance-oriented innovation rather than pursue meaningful technological upgrading that advances the sustainability transition. To better align green credit instruments with the objectives of high-quality, innovation-driven sustainable development, policymakers should further improve credit accessibility for non-SOEs and small firms, strengthen monitoring mechanisms that link financial incentives to patent quality rather than quantity and to verifiable sustainability outcomes, and refine institutional arrangements that encourage long-term technological breakthroughs and sustained environmental performance improvements. Such policy adjustments are essential for enhancing the effectiveness of green finance in supporting environmentally sustainable and innovation-led economic transformation and for ensuring that green credit contributes to carbon-neutral and sustainable development trajectories.

6.2 Recommendations

Based on the empirical findings, this study offers several policy and managerial recommendations to enhance the effectiveness of green credit in promoting high-quality and sustainability-oriented innovation.

First, policymakers should refine the design of green credit evaluation systems by shifting the focus from innovation quantity to innovation quality and verifiable sustainability outcomes. Rather than relying predominantly on patent counts, financial institutions should incorporate indicators related to patent novelty, technological impact, and actual environmental performance improvements, such as emissions reduction and energy efficiency gains. This adjustment would help mitigate firms' incentives to engage in symbolic green innovation and reduce the risk of greenwashing behavior under sustainability-oriented financial regulation.

Second, differentiated green credit support mechanisms should be strengthened to address firm heterogeneity. The results indicate that non-SOEs and small firms face tighter financing constraints that hinder their ability to pursue substantive green innovation. Policymakers should therefore improve inclusive access to green finance, for example through credit guarantees, interest subsidies, or risk-sharing arrangements, to support these firms' participation in long-term sustainable technological upgrading. Such measures would enhance the equity and effectiveness of green finance in fostering economy-wide sustainability transitions.

Third, financial regulators and banks should enhance post-credit monitoring and information disclosure mechanisms to ensure that green credit contributes to sustained environmental and sustainability performance, rather than short-term compliance. Link-ing loan conditions and renewal decisions to firms' long-term innovation trajectories and environmental outcomes can strengthen the credibility of green finance and better align financial incentives with carbon neutrality and sustainable development goals.

6.3 Limitations and future research

Despite its contributions, this study has several limitations that point to promising avenues for future research. First, this study focuses on green patents as proxies for green innovation and sustainability-oriented technological activity. Although patent-based measures are widely used, they may not fully capture firms' actual environmental performance or broader sustainability impacts. Future research could integrate firm-level emissions data, energy consumption indicators, or ESG performance metrics to provide a more comprehensive assessment of sustainability outcomes.

Second, while the analysis identifies significant heterogeneity across firm ownership and size, other sources of heterogeneity remain unexplored. Future studies could examine how regional institutional environments, industry-specific characteristics, or differences in local financial development shape the sustainability effects of green credit policies.

Third, this study examines the Green Credit Policy within the Chinese institutional context. Although China provides a valuable setting for studying green finance and sustainability-oriented regulation, the findings may not be fully generalizable to other economies with different financial systems or regulatory frameworks. Comparative studies across countries or regions would help assess the external validity of green credit as a tool for promoting sustainable development.

Finally, future research could explore the dynamic and long-term effects of green credit policies on firms' innovation trajectories and environmental performance. Understanding whether symbolic innovation evolves into substantive innovation over time would offer deeper insights into the long-run sustainability implications of green finance.

Funding

No

Conflict of Interests

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

Reference

- [1] Lv, C., Bian, B., Lee, C. C., & He, Z. (2021). Regional gap and the trend of green finance development in China. *Energy Economics*, 102, 105476.
- [2] Li, Y., Jia, Z., Qin, C., Li, Z., & Ying, J. (2025). Analysis of recycling and emission reduction models under carbon taxes and government subsidies. *International Review of Economics & Finance*, 99, 104052.
- [3] Zhang, K., Li, Y., Qi, Y., & Shao, S. (2021). Can green credit policy improve environmental quality? Evidence from China. *Journal of Environmental Management*, 298, 113445.
- [4] Acemoglu, D., Aghion, P., Bursztyn, L., & Hémous, D. (2012). The environment and directed technical change. *American economic review*, 102(1), 131-166.
- [5] Wen, H., Lee, C. C., & Zhou, F. (2021). Green credit policy, credit allocation efficiency and upgrade of energy-intensive enterprises. *Energy Economics*, 94, 105099.
- [6] Ma, Y., & Chen, D. (2025). How green credit policies drive green innovation: evidence from a natural experiment in China. *Journal of Economic Behavior & Organization*, 239, 107265.
- [7] Yuan, X., Shang, L., & Xu, J. (2024). Green Financial Policy, Resource Allocation and Corporate Environmental Responsibility. *Sustainability*, 16(15), 6273.
- [8] He, J., Xue, H., Yang, W., Zhong, Y., & Fan, B. (2025). Green credit policy and corporate green innovation. *International Review of Economics & Finance*, 99, 104031.
- [9] Wang, Y., Lei, X., Long, R., & Zhao, J. (2020). Green credit, financial constraint, and capital investment: Evidence from China's energy-intensive enterprises. *Environmental Management*, 66(6), 1059-1071.

- [10] Sun, W., Mei, J., Chen, K., & Li, Z. (2025). Biodiversity risk and value creation in emerging markets. *Finance Research Letters*, 108069.
- [11] Huang, J., & Ma, L. (2024). Substantive green innovation or symbolic green innovation: The impact of fintech on corporate green innovation. *Finance Research Letters*, 63, 105265.
- [12] Zhang, Z., Hou, Y., Li, Z., & Li, M. (2024). From symbolic to substantive green innovation: How does ESG ratings optimize corporate green innovation structure. *Finance Research Letters*, 63, 105401.
- [13] Li, X., & Li, H. (2025). Bright minds or green Façades? Superstar inventors and substantive vs. strategic green innovation. *Total Quality Management & Business Excellence*, 1-33.
- [14] Song, C., & Ma, W. (2025). ESG and green innovation: nonlinear moderation of public attention. *Humanities and Social Sciences Communications*, 12(1), 1-14.
- [15] Wang, F., Yang, S., Reisner, A., & Liu, N. (2019). Does green credit policy work in China? The correlation between green credit and corporate environmental information disclosure quality. *Sustainability*, 11(3), 733.
- [16] Shao, C., Wei, J., & Liu, C. (2021). Empirical analysis of the influence of green credit on the industrial structure: A case study of China. *Sustainability*, 13(11), 5997.
- [17] Li, Y., Ding, T., & Zhu, W. (2022). Can green credit contribute to sustainable economic growth? An empirical study from China. *Sustainability*, 14(11), 6661.
- [18] Chen, Z., Zhang, Y., Wang, H., Ouyang, X., & Xie, Y. (2022). Can green credit policy promote low-carbon technology innovation?. *Journal of cleaner production*, 359, 132061.
- [19] Ma, Y., Lu, L., Cui, J., & Shi, X. (2024). Can green credit policy stimulate firms' green investments?. *International Review of Economics & Finance*, 91, 123-137.
- [20] Wang, H., Qi, S., Zhou, C., Zhou, J., & Huang, X. (2022). Green credit policy, government behavior and green innovation quality of enterprises. *Journal of Cleaner Production*, 331, 129834.
- [21] Zhang, Z., Duan, H., Shan, S., Liu, Q., & Geng, W. (2022). The impact of green credit on the green innovation level of heavy-polluting enterprises—evidence from China. *International Journal of Environmental Research and Public Health*, 19(2), 650.
- [22] Porter, M. E., & Linde, C. V. D. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of economic perspectives*, 9(4), 97-118.
- [23] Jing, F., Muhamad, H., Said, R. M., & Daud, Z. M. (2024). Green Credit Policy and Firms' Green Total Factor Productivity: The Mediating Role of Financial Constraints. *Theoretical and Practical Research in Economic Fields*, 15(4), 871-884.
- [24] Kajejy, M. O., & Alzubi, A. B. (2026). Financial Constraints, Environmental Innovation and Sustainable Competitive Advantage: A Multi-Theoretical Cross-Country Analysis. *Sustainability*, 18(2), 716.
- [25] Shao, X., Gao, K., Wang, T., Zhang, Y., & Wei, Q. (2023). Does green credit promote firm environmental performance? A new perspective of economic growth target constraints. *Environmental Science and Pollution Research*, 30(50), 108617-108634.
- [26] Wang, D., Zhang, X., & Kwon, D. (2026). Environmental Regulation and Corporate Performance: Evidence from Chinese Heavily Polluting Firms. *Systems*, 14(4), 373.
- [27] Lai, Z., Lou, G., Zhang, T., & Fan, T. (2023). Financing and coordination strategies for a manufacturer with limited operating and green innovation capital: Bank credit financing versus supplier green investment. *Annals of Operations Research*, 331(1), 85-119.
- [28] Wang, T., Liu, X., & Wang, H. (2022). Green bonds, financing constraints, and green innovation. *Journal of Cleaner Production*, 381, 135134.
- [29] Li, W., Cheng, H., He, J., Song, Y., & Bu, H. (2024). The impacts of green credit policy on green innovation of high-polluting enterprises in China. *Finance Research Letters*, 62, 105167.
- [30] Fang, X., Liu, M., & Li, G. (2024). Can the green credit policy promote green innovation in enterprises? Empirical

- evidence from China. *Technological and Economic Development of Economy*, 30(4), 899-932.
- [31] Li, J., Zhang, C., Zhang, J., Mi, Z., Liu, Z., Gong, L., & Lu, G. (2023). Incentive or constraint? Comprehensive impacts of green credit policy on industrial energy intensity. *Environmental Science and Pollution Research*, 30(46), 103101-103118.
- [32] Ramanathan, R., He, Q., Black, A., Ghobadian, A., & Gallea, D. (2017). Environmental regulations, innovation and firm performance: A revisit of the Porter hypothesis. *Journal of Cleaner Production*, 155, 79-92.
- [33] Li, C., Bu, W., & Liu, S. (2025). The synergistic impact of green credit policy and government subsidies on green innovation of heavily polluting firms. *Discover Sustainability*, 6(1), 1080.
- [34] Wang, Z., Chen, J., & Xue, X. (2025). Assessing the efficacy of green credit policy in fostering green innovation in heavily polluting industries. *Clean Technologies and Environmental Policy*, 27(1), 309-325.
- [35] Bai, E., Wu, K., Zhu, H., Zhu, H., & Lu, Z. (2024). How does China's green credit policy affect the innovation of high-polluting enterprises? From the perspective of innovation quantity and quality. *Plos one*, 19(5), e0302789.
- [36] Su, X., Pan, C., Zhou, S., & Zhong, X. (2022). Threshold effect of green credit on firms' green technology innovation: Is environmental information disclosure important?. *Journal of Cleaner Production*, 380, 134945.
- [37] Guo, R., Zhang, Y., Chen, K., Wang, Y., & Ning, L. (2025). Heterogeneous impact of green finance instruments on firms' green innovation novelty: Policy mix or mess?. *Energy Economics*, 144, 108315.
- [38] Xia, X., Huang, T., & Zhang, S. (2023). The impact of intellectual property rights city policy on firm green innovation: A Quasi-Natural Experiment based on a Staggered DID Model. *Systems*, 11(4), 209.
- [39] Wang, M., Wurgler, J., & Zhang, H. (2026). Policy uncertainty reduces green innovation. *Journal of Financial Economics*, 175, 104189.
- [40] Lu, Y., Gao, Y., Zhang, Y., & Wang, J. (2022). Can the green finance policy force the green transformation of high-polluting enterprises? A quasi-natural experiment based on "Green Credit Guidelines". *Energy Economics*, 114, 106265.
- [41] Xiao, B., Guo, X., Guo, X., & Wang, J. (2024). How does green finance policy in China help reduce pollution emissions? Energy efficiency improvement or green innovation. *Journal of Cleaner Production*, 467, 142933.
- [42] Kim, S., & Li, Z. (2021). Understanding the impact of ESG practices in corporate finance. *Sustainability*, 13(7), 3746.
- [43] Liu, X., Ji, Q., & Yu, J. (2021). Sustainable development goals and firm carbon emissions: evidence from a quasi-natural experiment in China. *Energy Economics*, 103, 105627.
- [44] Liu, Q., & Qiu, L. D. (2016). Intermediate input imports and innovations: Evidence from Chinese firms' patent filings. *Journal of International Economics*, 103, 166-183.
- [45] Sun, G., Feng, Q., Wu, S., & Xia, C. (2025). The impact of China's green credit policy on the innovation of manufacturing enterprises. *Journal of Innovation & Knowledge (JIK)*, 10(3), 1-14.
- [46] Lu, Q., Deng, Y., Wang, X., & Wang, A. (2024). The impact of China's green credit policy on enterprise digital innovation: evidence from heavily-polluting Chinese listed companies. *China Finance Review International*, 14(1), 103-121.
- [47] Chang, K., Zeng, Y., Wang, W., & Wu, X. (2019). The effects of credit policy and financial constraints on tangible and research & development investment: Firm-level evidence from China's renewable energy industry. *Energy policy*, 130, 438-447.

Research on the Impact of Long Term Care Insurance on Family Medical Expenditure: Empirical Evidence Based on CHARLS Data from 2011 to 2020

Churui Yang¹, Rongxin Li^{2*}, Xunwen Zheng³

1. School of Humanities and Social Sciences (School of Public Administration), Beihang University, Beijing, 100191, China

2. School of Ecology and Environment, Renmin University of China, Beijing, 100872, China

3. Graduate School of Economics and Faculty of Economics, Kyoto University, Kyoto City, Japan

*Corresponding author: Rongxin Li, lirongxin7868@ruc.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 recent years, the long-term care insurance system has played an important role in the social security system under the background of population aging. Different from previous studies that focused on the improvement effect of policies on the health status of the elderly, the evaluation of policy effects, and the limitations of data lag, this study adopts the latest CHARLS data from 2011 to 2020 and uses a difference in differences model to explore the impact of long-term care insurance policies on family medical expenditures, and deeply analyzes the mediating and moderating effects. Research has found that long-term care insurance can significantly reduce household medical expenses; Mechanism analysis shows that this policy improves the physical health status of disabled elderly people by providing professional nursing services, thereby achieving a decrease in family medical expenses; The number of children has a negative moderating effect on the effectiveness of policies; Heterogeneity analysis shows that long-term care insurance has a more prominent effect on reducing medical expenses for families of chronic disease patients and highly educated families. Based on the above conclusions, it is recommended to adopt a gradual promotion strategy, gradually expand the coverage and guarantee objects of the pilot program, develop personalized service plans based on differences in disability level, economic status, etc., strengthen the professional training of nursing talents, and establish a multidimensional evaluation mechanism to regularly monitor the effectiveness of policy implementation and dynamically adjust the implementation plan.

Keywords: Long Term Care Insurance; Medical Expenses; Double Difference Model

Published: Apr 23, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1261>

1. Introduction

In 2015, China officially proposed to explore the establishment of a long-term care insurance system, which provides protection for medical expenses related to the basic living needs of disabled elderly people, in response to the care needs of disabled elderly people in the context of population aging, and reduces the expenses incurred by traditional small-scale families for caring for disabled elderly people (Cheng et al., 2023). Based on this, it can be seen that the long-term care insurance policy aims to alleviate the economic pressure of disabled elderly people in terms of home care, and to bear some of the main risks faced by individuals and society in long-term care.

As of January 2025, more than 180 million people in 49 national pilot cities in China have been insured, and over 2 million disabled individuals have received various forms of long-term care services such as institutional care, home care, and family care. Liu (2025) believes that the long-term care insurance program is an important component of the healthcare system aimed at supporting disabled elderly and disabled individuals and their families. In this process, long-term care insurance provides temporary care, some healthcare facilities, economic subsidies, etc., which reduces the original health and medical expenses of these families; Zhao (2025) believes that long-term care insurance has a significant inhibitory effect on the end-of-life medical expenses paid by elderly medical insurance. Long term care insurance reduces the probability of elderly people suffering from major diseases, improves their physical health status, and thus reduces the end-of-life medical expenses of the elderly; Ma (2019) believes that long-term care insurance saves outpatient and inpatient expenses for the elderly, effectively reducing medical insurance fund expenditures. More importantly, the medical expenses saved by long-term care insurance are not based on harming the health of middle-aged and elderly people. To some extent, this insurance has also played a positive role in improving the mental health status of middle-aged and elderly people, while also reducing the likelihood of their physical pain.

On the other hand, Zhu Minglai believes that long-term care insurance can significantly improve the daily activity ability and health level of disabled elderly people by increasing the quality of care for the elderly and replacing children's care; Zhou (2024) believes that long-term care insurance has a positive spillover effect on the employment of children in elderly families, reducing their financial vulnerability and thus improving their overall income situation. In this situation, Chen et al. (2024) believes that long-term care insurance can squeeze out family care, create employment opportunities, and significantly increase household income levels. After the increase in household income, Chen (2014) believes that , household health and medical expenses are related to household income, and health shocks are a direct factor in household asset behavior choices. Therefore, while long-term care insurance improves the health of the elderly, it also increases household income, and medical and health expenditures will actually increase due to the increase in household income.

Coe (2023) believes that the coverage of long-term care insurance in the United States reduces parents' perception of their children's willingness to take care of them in the future, and long-term care insurance has spillover effects on the economic behavior of family members; Hyunjong Song (2025) believes that the long-term care insurance policy introduced by South Korea on July 1, 2008, aims to alleviate the burden on families caring for the elderly in the context of rapid aging and social change. However, there are still shortcomings in fully meeting the end-of-life care, health, and medical needs of the elderly. It is necessary to reform the long-term care insurance system and incorporate end-of-life care into the existing long-term care insurance framework to improve service coverage and enhance the quality of care for the elderly. Patrizio (2024) believes that the aging population in Germany is severe, and the number of elderly people has been continuously increasing in recent years. The government has formulated long-term care insurance policies to meet the rapidly growing demand for nursing staff. Based on population changes, epidemiological trends, specific needs and supply of nursing services, and corresponding costs, predictive models are designed and used as a case study for long-term care insurance in Germany until 2050.

After the implementation of long-term care insurance policies, the health and financial situation of disabled elderly people have been significantly improved (Shu et al., 2022). In the protection system of long-term care insurance, the service items cover two dimensions: life assistance and professional nursing. Specifically, this system can significantly improve the quality of life of the elderly population by providing daily care services such as personal hygiene assistance and dietary management, combined with professional nursing interventions such as rehabilitation training and disease prevention. On the other hand, long-term care insurance policies can further share the responsibility of caring for children by providing more favorable care services and policy based welfare benefits for disabled elderly people, thereby increasing children's labor time, increasing family income, and reducing the burden of elderly care. However, due to the increase in household income, there will also be an increase in medical expenses and nursing needs, resulting in an overall increase in household health and hygiene expenditures. Therefore, overall, the relationship and interaction between long-term care insurance policies and household medical expenditures are still not clear enough, and further research is needed based on the survey data.

The relationship between long-term care insurance system and family medical expenses has become a widely discussed

topic in the academic community. Through reviewing existing literature, it is found that domestic and foreign research mainly focuses on the impact of this system on improving the health status of the elderly, changes in household consumption structure, and the transformation of intergenerational care models for disabled elderly. Research has been conducted on the impact of long-term care insurance on the health status, household consumption, and intergenerational care of disabled elderly. However, there is still a clear gap in specialized research on family medical expenditures, and there is no consensus on existing results. At the same time, the academic exploration of the influencing factors of family medical expenditure mainly focuses on traditional variables such as macroeconomic environment, family economic foundation, and the health status of the elderly, lacking in-depth examination of the social security system.

In existing literature on the evaluation of long-term care insurance, there is a common problem of generalization of research subjects. For example, existing studies often mix insured and uninsured populations for analysis, which may lead to bias in evaluating the actual effectiveness of policies. Therefore, this article mainly conducts in-depth research on the relationship and mechanism between long-term care insurance and family medical expenses.

2. Theoretical assumptions

This article mainly discusses the impact of the pilot system of long-term care insurance in China on family medical expenses. Long term care insurance provides care services and economic compensation for disabled elderly people, significantly reducing the care time for their children and reducing the use of medical resources, promoting the optimization of medical resources, and reducing family medical expenses (Zhu et al., 2023). Therefore, this article believes that long-term care insurance has a suppressive effect on family medical expenses and proposes the following hypothesis (Xie et al., 2024).

Assumption 1: Long term care insurance can significantly reduce household medical expenses.

In the process of reducing family medical expenses through long-term care insurance policies, the physical health status of elderly people plays a key mediating role (Zhu et al., 2024). Long term care insurance first affects the mediating variable of health status through resource investment, and then indirectly reduces the burden of family medical care through changes in health status. For example, for the elderly with chronic diseases such as hypertension and diabetes, the regular blood glucose and blood pressure monitoring and medication guidance covered by the long-term care insurance can effectively control the stability of the condition and reduce the cost of hospitalization due to complications. At this time, the stability of the health status directly mediates the inhibitory effect of policies on expenditure.

In addition, the mediating effect of physical health status is also reflected in preventive health management (Yue et al., 2021). If insurance policies include services such as health screening and early intervention, health risks can be identified and intervened in advance to avoid the transformation of health problems into disease states, thereby reducing medical expenses at the source. Based on this, this article proposes hypothesis 2:

Assumption 2: Long term care insurance affects household medical expenses through the health status of the elderly

Long term care insurance mainly targets the care of the elderly. In the process of exploring the impact of long-term care insurance on family medical expenses, elderly people with chronic diseases often require more medical resources and medical expenses (Xie et al., 2022). Their daily expenses are significantly higher than those of elderly people without chronic diseases; On the other hand, families with higher education levels have richer knowledge in health and hygiene, attach more importance to their own health conditions, and have a higher proportion of medical expenses in their living expenses. They will choose more expensive medical services and products; Families with higher education levels often have more social and economic resources, resulting in relatively higher medical expenses. Long term care insurance can significantly reduce the family medical expenses of elderly people by providing care and assistance to highly educated samples (Zhou et al., 2024). Based on this, hypotheses 3 and 4 are proposed:

Assumption 3: The long-term care insurance policy has a stronger effect on reducing medical expenses for families with chronic diseases.

Assumption 4: The long-term care insurance policy has a stronger effect on reducing medical expenses for highly educated families.

The main source of income for families of elderly people with disabilities is the income of their children. In families with

more children, the total income of the family is often higher and the financial situation is better (Li S et al., 2024). The expenses for caring for the elderly and medical care are also correspondingly higher. At the same time, the probability of elderly people choosing to give up treatment and care after becoming disabled is also lower. Children will spend more money to assist the elderly, and even if disabled elderly people have received care services during the long-term care insurance process, their children will continue to receive additional subsidies and medical expenses.

In families with fewer children, the family income is relatively low, and elderly people often do not choose to increase the proportion of medical expenses. They prefer to leave the money to their children, and their children will not provide too much medical care for the elderly. In the process of the elderly's death, there is a greater possibility of giving up treatment (Wang et al., 2021). After long-term care insurance provides corresponding care services, children do not need to continue to pay for healthcare related expenses for the elderly. Therefore, after the long-term care insurance takes effect, the medical expenses of such families often decrease significantly. Based on this, this article proposes hypothesis 5:

Assumption 5: The number of children will weaken the inhibitory effect of long-term care insurance on family medical expenses.

3. Data Explanation, Model Setting, and Variable Construction

3.1 Data Description

Referring to relevant articles on long-term care insurance research, this article uses survey data from five periods from 2011 to 2020 in the CHARLS database. According to the document “Guiding Opinions on Pilot Implementation of Long term Care Insurance System” released by the General Office of the Ministry of Human Resources and Social Security of China in 2016, 15 pilot cities were launched as the first batch of pilot areas for long-term care insurance. As the CHARLS database only contains survey data from 28 provinces and does not collect relevant respondent data from Shihezi, Nantong, and Changchun in Xinjiang, this article mainly evaluates the effectiveness of long-term care insurance policies in the remaining 12 cities. These 12 cities are set as the experimental group in the difference in differences model, and the remaining cities in CHARLS that have not implemented long-term care insurance are used as the control group.

3.2 Model Setting

The main research method used in this article is the double difference method, which mainly analyzes and evaluates the impact of long-term care insurance pilot policies on family medical expenditures. The constructed model is as follows:

$$Y_{ict} = \alpha_i + \beta_1 LTCI_{ct} + \delta_i control_{ict} + \lambda_c + \gamma_t + \sigma_{ict} \quad (1)$$

The subscripts i , c , and t in the model represent the individual being interviewed, the city where the sample family is located, and the time of the interview, respectively. The dependent variable Y_{ict} represents household medical expenses; The core explanatory variable $LTCI_{ct}$ of this article is the implementation status of the long-term care insurance pilot policy, indicating whether the sample studied in this article had implemented the long-term care insurance pilot policy during the time of receiving the CHARLS questionnaire survey. The LTCI variable value of the experimental group was 1 after 2016 and 0 before 2016, while the control group remained 0. Its estimated coefficient β_1 represents the impact of the long-term care insurance policy on family medical expenses; $control_{ict}$ refers to controlling variables, including gender, age, household registration type, marital status, and so on. In order to avoid omissions and biases caused by other factors, the model incorporates time fixed effects and region fixed effects, where λ_c and γ_t represent region and time fixed effects, respectively; σ_{ict} is a random perturbation term.

3.3 Variable Construction and Descriptive Statistics

3.3.1 Explained variable

The dependent variable in this article is household medical expenditure. Referring to the views of scholars Zhao Ming (2025) and Ma Chao (2019), “total outpatient consumption in the past month”, “number of outpatient visits in the past month”, “total hospitalization consumption in the past year”, and “number of hospitalizations in the past year” were also used as the dependent variables for benchmark regression and recorded as direct medical expenses, while transportation expenses, nutrition expenses, and family care expenses incurred due to medical treatment were recorded as indirect medical expenses; Family medical expenses include direct and indirect medical expenses, but do not include the portion already compensated

by medical insurance. Referring to the viewpoint of scholar Chen Jing (2020), the natural logarithm of the amount of family medical expenses is taken to comprehensively reflect the comprehensive level of family medical expenses.

3.3.2 Explanatory variable

The core explanatory variable is “long-term care insurance”, assigned by the specific pilot policies. Implementing “long-term care insurance” is designated as 1, otherwise the variable is 0. Since the implementation of China's long-term care insurance pilot system in 2016, there have been a total of 15 pilot cities nationwide. Due to the absence of relevant information on Changchun City, Nantong City, and Shihezi City in Xinjiang in the CHARLS questionnaire survey database, the main experimental groups in this paper are Chengde City in Hebei Province, Qiqihar City in Heilongjiang Province, Shanghai City, Suzhou City in Jiangsu Province, Ningbo City in Zhejiang Province, Anqing City in Anhui Province, Shangrao City in Jiangxi Province, Qingdao City in Shandong Province, Jingmen City in Hubei Province, Guangzhou City in Guangdong Province, Chongqing City, and Chengdu City in Sichuan Province. The remaining provinces and cities serve as the control group.

3.3.3 Control variable

The control variables include “gender”, “age”, “marital status”, “household registration type”, etc. Referring to the setting of control variables in Zhou Bowen's work (2024), we select the interviewer's gender, age, and other household demographic characteristics as control variables. Additionally, we control for household income, and in this paper, we log-transform the “household income” variable.

Table 1 Descriptive Statistics

Variable Name	Sample	Mean	standard deviation	minimum value	Maximum value
Family medical expenses	55478	6.21	3.29	0	13.99
Long term care insurance	55478	0.04	0.19	0	1
gender	55478	0.47	0.50	0	1
marital status	55478	0.86	0.35	0	1
household registration type	55478	0.77	0.42	0	1
Number of family members	55478	3.19	1.63	1	16
age	55478	60.68	10.49	11	120
household income	55478	37705.88	189481.60	0	39100000
Logarithm of household income	55478	9.20	2.49	0	17.48

4. Empirical Analysis

4.1 Benchmark Regression Analysis

This article conducts benchmark regression on the dependent variables such as family medical expenses, number of hospitalizations and outpatient visits, and out of pocket expenses. The regression process always controls for fixed effects of time and region, while controlling for a series of control variables such as gender, marital status, household registration type, number of family members, age, and family income, in order to more accurately identify the independent impact of long-term care insurance policies.

From the regression results, it can be seen that long-term care insurance has a negative impact on family medical expenses, monthly hospitalizations, annual outpatient visits, hospital out of pocket expenses, and outpatient out of pocket expenses. The coefficients are all negative and significant at the 1% level. This indicates that the implementation of long-term care insurance policy has a clear inhibitory effect on family medical related expenses and medical behavior, verifying hypothesis one.

By controlling for the influence of variables, long-term care insurance provides targeted care services and economic compensation for disabled elderly people, replacing some of the care needs that originally needed to be met through medical means, reducing the excessive use of medical resources, and alleviating the caregiving pressure and economic burden on family members. This effectively controls family medical costs from the perspectives of medical frequency and expenditure,

fully reflecting the policy value of long-term care insurance in optimizing medical resource allocation and alleviating family medical burden. It also provides empirical support for the promotion and improvement of the long-term care insurance system in the future.

Table 2 Benchmark regression results of the impact of long-term care insurance on household medical expenses

Variable	(1) Family medi- cal expenses	(2) Monthly hos- pitalizations	(3) Number of out- patient visits per year	(4) Self paid hospi- talization ex- penses	(5) Outpatient out of pocket expenses
Long term care insurance	-0.0135*** (-2.9556)	-0.1550*** (-3.4805)	-0.0468*** (-2.6596)	-0.1384*** (-3.0078)	-0.1972*** (-5.3180)
gender	-0.0099*** (-8.0771)	-0.1009*** (-8.3340)	-0.0128* (-1.6707)	0.0150 (0.7491)	-0.1257*** (-7.7757)
marital status	0.0006 (0.3238)	-0.0271 (-1.4461)	0.0070 (0.6231)	0.0475 (1.6138)	-0.0332 (-1.3984)
household registration type	-0.0049*** (-2.7902)	0.0206 (1.1887)	0.0383*** (4.0683)	-0.1951*** (-7.9242)	-0.0145 (-0.7326)
Number of family members	0.0019*** (4.4516)	-0.0000 (-0.0067)	0.0036 (1.4764)	0.0044 (0.6821)	0.0446*** (8.6528)
age	0.0014*** (20.8840)	0.0041*** (6.6367)	-0.0089*** (-22.0682)	0.0161*** (15.2367)	-0.0008 (-0.9499)
Logarithm of household income	-0.0004 (-1.2881)	0.0033 (1.2910)	0.0019 (1.1606)	-0.0268*** (-6.2470)	-0.0298*** (-8.6036)
constant term	0.0889*** (14.9343)	0.2044*** (3.7455)	0.2444*** (6.8571)	0.0730 (0.7836)	0.9771*** (13.0137)
Fixed time effect	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes
N	55478	55478	55478	55478	55478

Note: The values in parentheses are the statistical values of t; *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

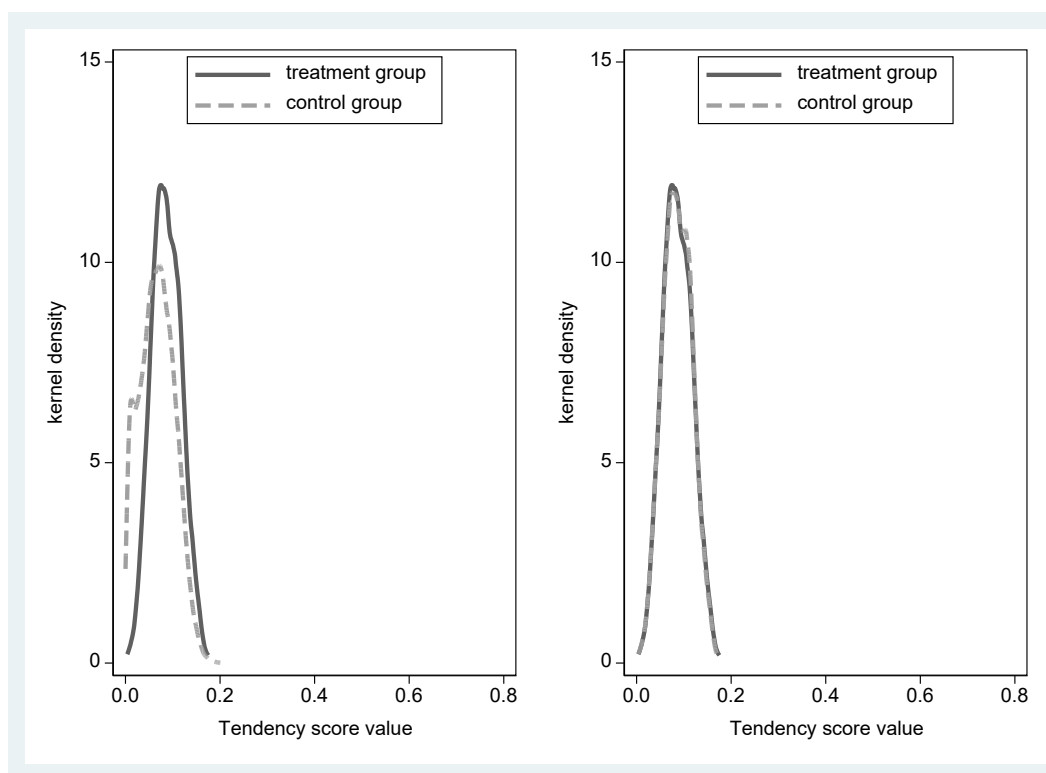
4.2 PSM-DID Inspection

The core function of PSM-DID is to optimize the comparability between the experimental group and the control group through a two-stage method: firstly, propensity score matching (PSM) constructs statistical weights based on observable covariates (such as household income, number of household members, etc.), selects control group samples with highly similar characteristics to the experimental group, and effectively alleviates baseline differences caused by non random sample allocation; Secondly, the double difference method (DID) eliminates the combined effects of time trends and individual fixed effects through differential processing, further separating the net effects of policy interventions.

After propensity score matching, the percentage deviation of all variables significantly decreased to within 10%, and the p-values of t-test were not significant ($p > 0.1$), indicating that the matching significantly reduced the covariate imbalance between the experimental group and the control group.

The results of the balance test indicate that a control group sample with highly similar characteristics to the experimental group was selected to achieve statistical balance between the two groups on key variables such as the number of chronic diseases, household registration type, gender, and income. This result not only satisfies the balance assumption of propensity score matching, but also lays a reliable foundation for subsequent DID analysis - by eliminating the influence of observable confounding factors before the policy, ensuring the accuracy of policy effect estimation.

Figure 1: Kernel Probability Density Map before and after Matching



Furthermore, the above figure shows the kernel probability density map before and after PSM matching, which is used to determine whether there is a difference in propensity score values between the two groups before and after matching. It can be seen that the two curves are closer after matching, indicating that the matching is effective. Based on the sample results matched by propensity score, the double difference method is used for regression.

Table 3 PSM-DID Regression Results

Family medical expenses	(1)	(2)
	benchmark regression	PSM-DID regression
Long term care insurance	-0.0135*** (-2.9556)	-0.0102*** (-5.3242)
gender	-0.0099*** (-8.0771)	-0.0077 (-1.0466)
marital status	0.0006 (0.3238)	-0.0005 (-0.1592)
household registration type	-0.0049*** (-2.7902)	-0.0089*** (-3.0891)
Number of family members	0.0019*** (4.4516)	0.0030*** (4.2076)
age	0.0014*** (20.8840)	0.0015*** (13.7297)
Logarithm of household income	-0.0004 (-1.2881)	-0.0003 (-0.7104)
constant term	0.0889*** (14.9343)	0.0832*** (8.9398)
Fixed time effect	Yes	Yes
Regional fixed effects	Yes	Yes
N	55478	32553

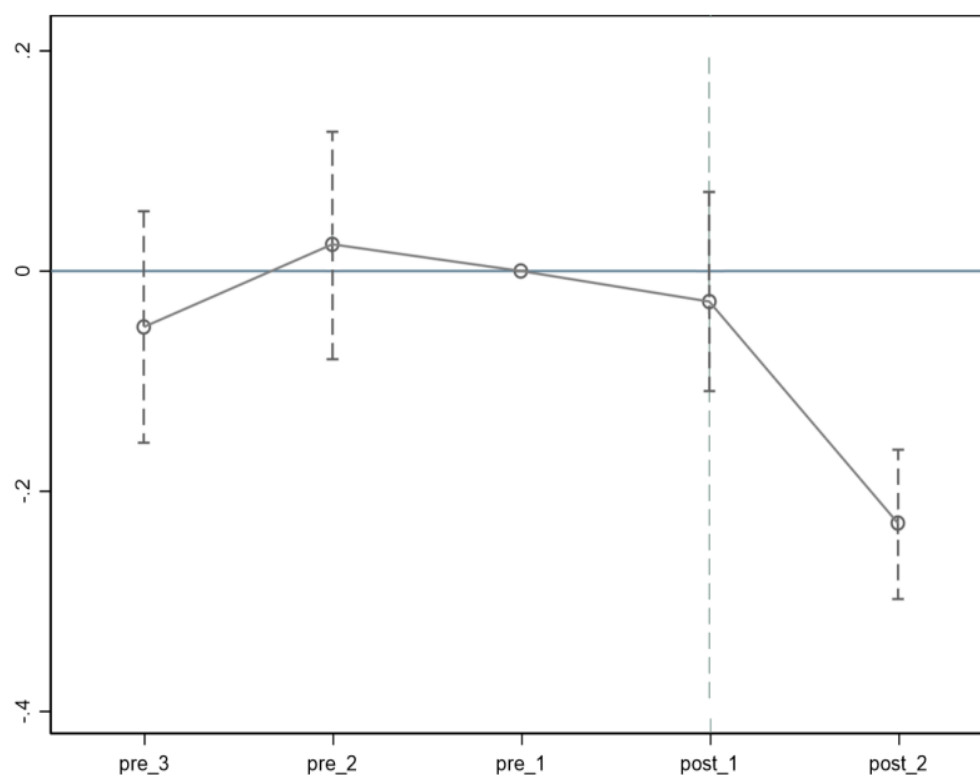
Note: The values in parentheses are the statistical values of t; *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

From the regression results of PSM-DID in the table above, it can be seen that long-term care insurance has a significant negative impact on household medical expenses at the 1% level, which fully proves that long-term care insurance policies can effectively reduce household medical expenses, thus verifying hypothesis one.

4.3 Parallel Trend Test

The parallel trend assumption is the core premise of the DID model. If there is a systematic difference in the trend of medical expenditure changes between the two groups before policy implementation, the subsequent estimation of policy effects may be mixed with unobservable confounding factors, leading to biased causal inference. Therefore, based on the China Health and Retirement Longitudinal Study (CHARLS) database, this study selected 12 pilot cities and non pilot areas for long-term care insurance, and constructed a time span to cover the data before and after the implementation of the policy for ten years (2011-2020). Using the policy launch year of 2016 as the base year, the pre policy period was divided into three periods (pre-1:2011, pre-2:2013, pre-3:2015), and the post policy period was divided into two periods (post-1:2018, post-2:2020). The dynamic effects model was used to capture the differences and evolution between the experimental group and the control group in each period. In the specific operation, the individuals in the pilot cities before policy implementation are first defined as the “virtual experimental group”, and individuals in non pilot areas are used as the control group. The model is used to estimate the interaction coefficients between the virtual variables (pre_1 to post_2) in each period and the virtual variables in the experimental group. These coefficients directly reflect the differences in medical expenditures between the experimental group and the control group before policy implementation, as well as the dynamic effects after policy implementation.

Figure 2: Parallel Trend Test Results



In order to visually present the test results, this study plotted the dynamic changes in coefficient values before and after the policy period (as shown in the above figure). If the coefficient before the policy is not statistically significant, it indicates that there is no systematic difference in trend between the experimental group and the control group before the policy, which satisfies the parallel trend hypothesis.

The empirical results show that the confidence intervals of the regression coefficients before policy implementation intersect with 0, and there is no significant difference in the trend of household medical expenditure between the experimental group and the control group in 2011, 2013, and 2015; After the implementation of the policy, the coefficient reached its maximum in the second phase, and the reduction effect of surface long-term protection insurance has a certain lag. This result strongly

supports the parallel trend hypothesis and provides a reliable basis for subsequent analysis of policy effects.

5. Mechanism of Action and Heterogeneity Analysis

5.1 Mechanism of Action

On the basis of the benchmark regression model 3.1, a two-stage regression model is further constructed to explore the transmission mechanism of long-term care insurance affecting family medical expenses through the health status of the elderly. The model is set as follows:

$$Health_{ct} = \alpha_0 + \alpha_1 LTCI_{ct} + \delta_1 control_{ict} + \lambda_c + \gamma_t + \sigma_{ict} \quad (2)$$

$$Y_{ict} = \beta_0 + \beta_1 LTCI_{ct} + \beta_2 Health_{ct} + \delta_1 control_{ict} + \lambda_c + \gamma_t + \sigma_{ict} \quad (3)$$

Among them, $Health_{ct}$ represents the mediating variable of physical health status. Referring to Xue (2025), the value of physical health status ranges from 0 to 33 points, with higher scores indicating better physical health status for the elderly; The parameter α_1 represents the impact of long-term care insurance on mediating variables; β_1 represents the direct effect of long-term care insurance on household medical expenses; $\alpha_1 \times \beta_2$ represents the indirect effect of long-term care insurance on household medical expenses.

Table 4 Regression Results of Mediating Effects

Explained variable	(1)	(2)	(3)
	Family medical expenses	health condition	Family medical expenses
Long term care insurance	-0.0135*** (-2.9556)	0.4172** (2.0680)	-0.0010 (-0.8043)
health condition			-0.0111*** (-6.1558)
control variable	Control	Control	Control
Fixed year effect	Control	Control	Control
Urban fixed effects	Control	Control	Control
N	55478	55478	55478

Note: The values in parentheses are the statistical values of t; *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

In the mediation regression results of the table above, model (2) shows that the implementation of long-term care insurance policy has a significant positive impact on health status at the 1% level, indicating that the health level of insured individuals is significantly improved compared to the uninsured group. The core mechanism of this effect is that long-term care insurance significantly improves the physical health status of disabled elderly people by covering professional nursing services such as home rehabilitation training and chronic disease management support.

After introducing health status as a mediator variable, the direct effect of long-term care insurance in model (3) decreased and was not significant; The coefficient of health status is significantly negative, indicating that long-term care insurance policies indirectly suppress medical expenditures by improving health, which verifies hypothesis 2.

5.2 Heterogeneity Analysis

5.2.1 Heterogeneity test of chronic diseases

Due to the long course of illness, multiple complications, and complex nursing needs, elderly people with chronic diseases often face higher consumption of medical resources, manifested in frequent outpatient follow-up, long-term medication expenses, professional nursing service fees, and the cost of purchasing assistive devices or home environment renovation caused by disability. These factors collectively push up the baseline level of family medical expenses (Ke J et al., 2025). Long term care insurance can theoretically alleviate family financial pressure by covering part of nursing expenses, providing home care subsidies, or reimbursing institutional care. However, its policy effect may show significant heterogeneity due to differences in chronic disease types, disease severity, and family resource endowments.

The study divided the total sample into two groups based on chronic disease status, with 42944 samples having chronic diseases and 12003 samples not having chronic diseases. According to the formula, the sample was regressed and heterogeneity analysis was conducted. The regression results are shown in the following table.

Table 5 Heterogeneity Analysis of Chronic Diseases

Family medical expenses	(1)	(2)
	Not suffering from chronic diseases	Suffering from chronic diseases
Long term care insurance	-0.0042 [*] (-1.9550)	-0.0189 ^{***} (-2.8501)
control variable	Control	Control
Fixed year effect	Control	Control
Urban fixed effects	Control	Control
<i>N</i>	12003	42944

Note: The values in parentheses are the statistical values of *t*; *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

The regression results showed that long-term care insurance significantly reduced household medical expenses in both groups, but there were significant differences in the strength of the effect. In samples without chronic diseases, the coefficient of long-term care insurance is only significant at the 10% level, and the coefficient is relatively small; Among the samples with chronic diseases, the policy significantly reduced household medical expenses, confirming that the policy dividends for families with chronic diseases are more significant. Therefore, it indicates that long-term care insurance has a stronger effect on samples with chronic diseases, significantly reducing their family medical expenses, thus verifying hypothesis 3.

5.2.2 Heterogeneity analysis of educational level

Education level, as a core proxy variable of social and economic status, shapes family healthcare decisions through a dual path of resource acquisition ability and health literacy: higher education groups usually have higher income levels, broader social networks, and better information acquisition channels, and can prioritize high-quality medical services and preventive health management, thus forming a “resource health” positive cycle; At the same time, its strong learning ability and health knowledge reserve enable it to identify disease risks earlier and optimize expenditure structure more accurately by utilizing insurance terms, such as selecting the most cost-effective service by comparing reimbursement ratios of different nursing institutions, or using policy allowed preventive care projects to reduce later treatment costs (Zhang et al., 2025). On the contrary, the low education group is limited by economic capital and health awareness, and their medical behavior relies more on passive treatment rather than active prevention. They also have a lower understanding and utilization efficiency of insurance policies, which may lead to insufficient release of policy dividends. From this, it can be concluded that long-term care insurance may have differences in medical expenses for families with different levels of education.

Table 6 Heterogeneity Analysis of Educational Level

Family medical expenses	(1)	(2)
	Primary school education or below	Primary school education or above
Long term care insurance	-0.0092 (-1.4066)	-0.0169 ^{***} (-2.6188)
control variable	Control	Control
Fixed year effect	Control	Control
Urban fixed effects	Control	Control
<i>N</i>	24833	30667

Note: The values in parentheses are the statistical values of t; *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

As shown in the table above, long-term care insurance has a significant inhibitory effect on both types of household medical expenditures, but the strength of the effect increases with the level of education. In the sample of primary school and below education, the policy effect is not significant; However, the coefficient of samples with primary school education or above is significantly negative, indicating that the long-term care insurance policy has a stronger effect on reducing household medical expenses for highly educated samples, which verifies hypothesis 4.

6. Further analysis

This study is based on family decision-making theory and health economics models, and explores in depth the moderating effect of the number of children on the long-term care insurance policy. On a theoretical level, medical decision-making in families with multiple children exhibits collective bargaining characteristics, especially in the care of elderly people with chronic diseases or disabilities. An increase in the number of children may affect medical expenses through two pathways: on the one hand, the participation of multiple people in decision-making may prolong the treatment period or choose higher cost treatment plans (such as intensive care or overseas medical treatment), forming a “decision-making inertia” that drives up expenses; On the other hand, economic sharing and emotional support among children may alleviate the financial pressure of a single decision maker, reduce the probability of giving up treatment, and indirectly maintain the rigid characteristics of medical expenditures.

As an external intervention tool, the policy effect of long-term care insurance may vary due to differences in the number of children - in families with fewer children, the economic compensation provided by insurance can directly replace the direct care cost of children and reduce medical expenses; In families with multiple children, insurance reimbursement may be seen as an additional resource that stimulates more active treatment investment, but instead weakens the effectiveness of policy cuts. Therefore, this article investigates whether there is a moderating effect of the number of children, and the results are shown in the table below.

Table 7 Regression Results of the Moderating Effect of the Number of Children

Family medical expenses	(1)	(2)
	Excluding interactive items	Including interactive items
Long term care insurance	-0.0179*** (-2.6375)	-0.0142*** (-3.1055)
Interaction term with the number of children		0.0017*** (2.8766)
number of children	0.0018*** (3.0280)	0.0019 (0.8192)
control variable	Control	Control
Fixed year effect	Control	Control
Urban fixed effects	Control	Control
<i>N</i>	21975	21975

Note: The values in parentheses are the statistical values of t; *, **, *** represent significance levels of 10%, 5%, and 1%, respectively.

This study systematically tested the moderating effect of the number of children on the impact of long-term care insurance policy on family medical expenditure by constructing a moderating effect model (Model 2). The regression results showed that the interaction item was significantly positive, and this result remained stable after controlling family income, the prevalence of chronic diseases, urban and rural registered residence, and the region time fixed effect. It verified the negative

moderating mechanism of the number of children on the reduction of family medical expenditure by long-term care insurance, and verified Hypothesis 5.

Its internal logic can be decomposed into three paths: firstly, the substitution effect of family care resources. Multi child families form informal nursing networks through intergenerational division of labor, such as shift based home care and shared responsibility for medical decision-making, which can partially replace the professional services provided by long-term care insurance (such as home care subsidies, institutional care reimbursements, etc.), reduce the economic dependence of families on insurance, and thus weaken the direct effect of policies on reducing expenditures; Secondly, the ethical inertia of medical decision-making is deeply influenced by the concept of “filial piety and longevity” in traditional Chinese filial piety culture. Even in cases where medical treatment is deemed ineffective, families with multiple children tend to maintain active treatment through collective bargaining (such as continuous use of intensive care, high priced targeted drugs, or overseas medical treatment) to avoid bearing the social stigma of “unfilial”. This decision-making rigidity causes insurance covered nursing fees to be reallocated to high cost treatment projects, forming an expenditure cycle of “policy compensation treatment upgrade”; Thirdly, the psychological effect of risk diversification is that families with multiple children are less sensitive to financial pressure due to the sharing of economic responsibilities (such as medical expenses and loss of work) and emotional support, and tend to view long-term care insurance as a supplementary resource rather than a necessary guarantee, resulting in lower than expected insurance utilization efficiency.

7. Conclusion and Policy Suggestions

This study is based on CHARLS 2011-2020 panel data and uses a double difference method to systematically evaluate the impact of long-term care insurance policies on household medical expenditures and their pathways of action. The main research conclusions of this article are as follows:

Firstly, the long-term care insurance policy significantly reduces household medical expenses, and after multiple robustness tests, the research conclusion still holds true. Long term care insurance directly replaces the out of pocket medical costs of families by providing professional nursing services for disabled elderly people. At the same time, it indirectly reduces the incidence of complications and subsequent treatment needs by improving the health status of insured individuals. The policy effect of improving health status indicates that health improvement is an important transmission path for policy reduction. Secondly, the study revealed significant heterogeneity in the effectiveness of long-term care insurance policies. The cost reduction effect of chronic disease families is significantly higher than that of non chronic disease families, as their rigid medical needs are higher, and the nursing services covered by the policy can directly replace high treatment costs; The moderating effect of education level is significant, and the policy dividends for families with higher education are more prominent than those with education below high school, reflecting the advantages of the higher education group in understanding insurance terms, resource integration, and preventive health management. Thirdly, the moderating effect of the number of children weakens the policy effect through “alternative care resources” and “ethical decision-making inertia”, indicating that collective treatment decisions under traditional filial piety culture offset the economic compensation function of insurance.

Based on the above research results, this article proposes relevant policy recommendations:

One suggestion is to optimize the differentiated design of long-term care insurance: for families with chronic diseases, expand the reimbursement scope of nursing services (such as psychological rehabilitation and remote medical care), and establish a dynamic subsidy mechanism to match disease progression; For non chronic disease families, strengthen the coverage of preventive services (such as health screening and early intervention for chronic diseases), guide them to shift from “passive treatment” to “active health management”, and thus improve the overall efficiency of policies. The research results not only verify the cost control function of long-term care insurance, but also provide empirical evidence for the construction of a precise medical security system, helping to achieve the goal of fair and sustainable access to medical resources under the “Healthy China” strategy.

Secondly, the research findings call for the development of education sensitive long-term care insurance optimization strategies: for higher education groups, personalized insurance products should be developed (such as high-end nursing

service packages, cross-border medical reimbursements), and digital tools (such as AI insurance consultants) should be used to improve service accuracy; For the low education group, it is necessary to simplify the language of policy promotion, popularize insurance knowledge through community lectures, graphic manuals, and other forms, while increasing the reimbursement ratio for basic nursing services (such as family doctor contracts and chronic disease management), and lowering their medical expenditure threshold (Yin, M., & Hu, M. Y. , 2016).

Thirdly, targeted optimization of institutional design: for families with multiple children, a “nursing treatment” expenditure classification reimbursement mechanism should be established, and the reimbursement ratio for basic nursing services (such as home care and rehabilitation training) should be increased to 90%. For high-end treatment projects (such as ICU maintenance and experimental therapy), an annual payment limit (such as 50000 yuan) should be set, guiding resources to tilt towards nursing fields with higher cost-effectiveness; At the same time, the “Family Decision Support Program” is being implemented, which involves professional social workers intervening in medical consultations for families with multiple children, providing end-of-life care knowledge training and ethical conflict resolution, and alleviating irrational treatment investment.

For families with fewer children, it is necessary to strengthen the accessibility of insurance, such as simplifying the reimbursement process in different places, expanding the network coverage of nursing institutions, and exploring bundled products of “insurance+commercial supplements” to compensate for the shortage of nursing manpower. The limitations of the study include the heterogeneity of the impact of children's living distance and economic contribution, and the possibility of measurement errors in the proxy variable of filial piety concept. In the future, dynamic decision-making details can be captured through follow-up surveys and in-depth interviews.

Fourthly, at present, the number of pilot cities for long-term care insurance in China is relatively limited, and the regional distribution is concentrated in the economically developed eastern regions, resulting in significant regional imbalances. In the future, on the basis of existing pilot projects, priority should be given to including cities in the central and western regions with high aging rates and relatively scarce medical resources, gradually building a nationwide pilot network. At the same time, the current policy mainly covers the urban employee medical insurance insured group, and rural residents have been excluded from the security system for a long time. Compared to cities, the aging situation in rural areas is more severe, and the shortage of elderly care resources is prominent. Therefore, it is necessary to break down the dual barriers between urban and rural areas, extend the pilot scope to counties and rural areas, and focus on covering areas with high risk of disability and urgent care needs. In addition, a dynamic adjustment mechanism should be established to timely include individuals with moderate to mild disabilities, cognitive impairment patients, etc. in the scope of protection, forming a hierarchical and classified protection system.

Fifthly, in response to the diverse needs of the elderly population, it is necessary to abandon the “one size fits all” service model and establish a personalized care system that accurately identifies and dynamically responds. Specifically, for the home-based elderly care group, home care services (such as wound treatment and rehabilitation training) and subsidies for home aging adaptation (such as the installation of anti slip facilities) should be strengthened; For institutional elderly care recipients, focus on improving specialized medical care (such as postoperative rehabilitation, chronic disease management) and psychological support services; Exploring low-cost community mutual aid models (such as neighborhood care networks) for empty nest elderly in rural areas. At the same time, beneficiaries are given the right to choose independently, allowing for flexible allocation between cash subsidies and service payments, and encouraging family members to participate in care training (such as basic nursing skills courses) through policy support, forming a “family community institution” collaborative service ecosystem.

Sixthly, the key to improving the quality of nursing services lies in optimizing the professional abilities of practitioners. At the higher education level, elderly care majors should be added to medical colleges, offering interdisciplinary courses that integrate pathology, rehabilitation medicine, and psychology; In the field of vocational education, we will improve the short-term skills training system (such as disability assessment and emergency operations), and implement a dual track assessment and certification system of “theory+practice”. Establish regional training bases to simulate real care scenarios (such as home

care and institutional collaboration) and strengthen the emergency response capabilities of practitioners. Implement an annual mandatory continuing education program for employees, focusing on modules such as mental health intervention and cross-cultural communication skills.

In addition, it is necessary to systematically improve the nursing environment: firstly, improve the salary guarantee mechanism and implement tiered subsidies based on service difficulty and working hours; Secondly, establish a system for monitoring labor intensity and implementing job rotation and compensatory leave to avoid occupational burnout; The third is to enhance the social recognition of the nursing profession and attract high-quality talents to join the field of elderly care through media promotion and social honor selection. Through the above measures, a virtuous cycle of “talent supply service quality policy effectiveness” can be achieved.

Funding

No

Conflict of Interests

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

Reference

- [1] Cheng, Q., Lu, J. H., & Zheng, B. F. (2023). The impact of long-term care insurance system on middle-aged and elderly consumption: Evidence from CHARLS panel data. *Chinese Journal of Population Science*, 37(3), 82–96.
- [2] Liu, Y., Xu, W., Yang, Z., et al. (2025). Caregiving experiences of family caregivers of disabled middle-aged and older adults enrolled in long-term care insurance: A qualitative study. *BMC Nursing*, 24(1), 67. <https://doi.org/10.1186/s12912-025-01281-9>
- [3] Zhao, M., Zhao, Z. T., She, Z. X., et al. (2025). Long-term care insurance, end-of-life medical expenses, and value-based healthcare effects. *Insurance Studies*, (1), 68–80.
- [4] Ma, C., Yu, Q. W., Song, Z., et al. (2019). Long-term care insurance, medical cost control, and value-based healthcare. *China Industrial Economics*, (12), 42–59.
- [5] Zhu, M. L., & He, M. (2021). Does long-term care insurance crowd out family care? An empirical analysis based on CHARLS data from 2011 to 2018. *Insurance Studies*, (12), 21–38.
- [6] Zhou, B. W., & Zang, X. H. (2024). The impact of long-term care insurance on the financial vulnerability of elderly households: An empirical analysis based on CHARLS data. *Economic Perspectives*, (12), 111–127.
- [7] Chen, F., & Wang, R. T. (2024). Common prosperity effects of the long-term care insurance system: From the perspective of care decisions and risk sharing. *The Journal of World Economy*, 47(4), 154–183.
- [8] Chen, Q., & Liu, W. (2014). The impact of health expenditure on residents' asset allocation behavior: A discussion based on the homogeneity and heterogeneity debate. *Shanghai Journal of Economics*, (6), 111–118.
- [9] Costa, B. N., Shah, G. G., & Von Hinke, H. C. (2023). Family spillovers and long-term care insurance. *Journal of Health Economics*, 90, 102781. <https://doi.org/10.1016/j.jhealeco.2023.102781>
- [10] Kim, S., & Yoo, J. (2025). Effects of copayment reduction in long-term care insurance on medical utilization in South Korea. *Journal of Aging & Social Policy*, 37(3), 1–21. <https://doi.org/10.1080/08959420.2025.2321245>
- [11] Vanella, P., Wilke, B. C., & Heß, M. (2024). Long-term care in Germany in the context of the demographic transition—An outlook for the expenses of long-term care insurance through 2050. *Econometrics*, 12(4), 28. <https://doi.org/10.3390/econometrics12040028>
- [12] Huang, J. M., & Yu, X. L. (2025). Can long-term care insurance promote household consumption? A micro-perspective based on relative deprivation. *Collected Essays on Finance and Economics*, 1(13). Advance online publication.
- [13] Shu, Z., & Han, Y. (2022). The impact of long-term care insurance on intergenerational support among families of disabled elderly. *Population and Development*, 28(4), 143–154, 117.
- [14] Zhu, Z. Y. (2023). The impact of long-term care insurance on elderly family care. *Chinese Journal of Population Science*, 37(3), 97–114.

- [15] Xie, M. M., Yang, J., & Xie, X. (2024). The impact of long-term care insurance on catastrophic health expenditure among the elderly. *Population and Development*, 30(4), 156–166.
- [16] Zhu, J., Li, W., & Li, Z. F. (2024). The impact of long-term care insurance on rural household consumption levels: An empirical analysis based on CHARLS data. *The Theory and Practice of Finance and Economics*, 45(2), 104–111.
- [17] Yue, W., Wang, X., & Zhang, Q. (2021). Health risk, medical insurance, and household financial vulnerability. *China Industrial Economics*, (10), 175–192.
- [18] Xie, Y. F., & Feng, J. (2022). Has long-term care insurance narrowed the health gap among disabled elderly?. *Insurance Studies*, (10), 19–33.
- [19] Li, S., Zhang, L., & Fang, Y. (2024). Does social support alleviate the caregiving burden of adult children? Evidence from Chinese long-term care insurance pilot program. *Journal of Aging & Social Policy*, 37(3), 11–16. <https://doi.org/10.1080/08959420.2024.2321244>
- [20] Wang, Z., & Feng, J. (2021). 、 The substitution effect of long-term care insurance on medical expenses and a comparison of different compensation models、 . *China Economic Quarterly*, 21(2), 557–576.
- [21] Chen, J., & Sun, P. (2020). The impact of population aging on household health expenditure in China: An analysis based on microdata. *Productivity Research*, (11), 114–118.
- [22] Ke, J., & Sun, F. (2025). Long-term care insurance and health inequality: Evidence from China. *The International Journal of Health Planning and Management*, 40(3), 594–606. <https://doi.org/10.1002/hpm.3512>
- [23] Zhang, K., Liu, Y., & Hu, H. (2024). Multidimensional poverty and disability of older adults in China: Will long-term care insurance make a difference? *Applied Research in Quality of Life*, 19(6), 1–24. <https://doi.org/10.1007/s11482-024-10281-9>
- [24] Sun, Z., Li, Y., & Gao, S. (2024). Residents' cognition, attitudes, and intentions to participate in long-term care insurance: Moderating effect of policy support. *Behavioral Sciences*, 14(10), 895. <https://doi.org/10.3390/bs14100895>
- [25] Yin, M., & Hu, M. Y. (2016). Government health expenditure, economic development, and residents' health expenditure: An empirical test based on a panel threshold model. *Taxation and Economy*, (6), 30–36.
- [26] Xue, H. Y., & Zhang, Y. G. (2025). The impact of long-term care insurance on the health of the elderly: An empirical analysis based on CHARLS data. *Journal of Guizhou University of Finance and Economics*, (3), 20–30.

The Spillover Effect of Financial Agglomeration on Urbanization: An Empirical Study in Southwestern China

Xinyu Liu*

Sichuan Tianyi College, Mianzhu, Sichuan, 618200, China

*Corresponding author: Xinyu Liu, 1661161029@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: To investigate the spillover effect of financial agglomeration, this study selected 46 prefecture-level cities in Southwest China from 2011 to 2020 as research subjects. The research applies Moran's Index correlation test, using spatial econometric models for association analysis. Results show financial agglomeration has a pronounced spatial correlation with urbanization and the spillover effect of financial agglomeration enhances the development of urbanization. This study provides empirical support for policymakers in Southwest China to implement sustainable development goals for finance and urbanization.

Keywords: Spillover Effect; Financial Agglomeration; Urbanization; Southwestern China

Published: Apr 24, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1287>

1. Introduction

Financial agglomeration generally refers to the concentration of financial institutions and financial services within a specific area. This concentration can generate economies of scale, external economic effects, and knowledge spillovers, thereby improving the efficiency and competitiveness of regional financial services.^[1] The concept and phenomenon of financial agglomeration have been examined and developed by numerous scholars over time.^[2] Existing studies have identified a statistically significant spatial correlation between financial agglomeration and a range of economic activities.^[3-5] In particular, previous research has primarily focused on the effects of financial agglomeration on economic growth, employment, innovation, real estate, and green development. As a low-carbon industry, the financial sector plays an important role in urban sustainable development and clean production. Financial agglomeration is therefore regarded as an important driver of regional development, as it improves the allocation efficiency of financial resources, reduces transaction costs, and promotes urbanization and sustainable development. In developing countries and regions, financial agglomeration can also facilitate sustainable urbanization by encouraging industrial upgrading and environmental protection. Moreover, its influence extends beyond local urban development and may produce spillover effects in surrounding areas.^[8]

Urbanization is commonly understood as the process through which population shifts from rural areas to urban areas. However, it is a multidimensional and highly complex phenomenon. In most countries, the urbanization process can be broadly divided into five stages: economic growth and development, demographic transition, social transformation, the restructuring and expansion of urban space, and urban shrinkage planning. Current research on the determinants of urbanization mainly concentrates on economic, policy, social, geographical, cultural, and technological factors. Urbanization

affects economic growth through several important channels. First, cities play a central role in the economic and social structures of both developed and developing countries by providing access to education, employment, and healthcare. Second, urbanization promotes the concentration of population and enterprises, which generates economies of scale, lowers transaction costs, and creates a favorable environment for business development. This process also makes it easier for firms to obtain financing and access larger local markets. Urban enterprises are often more profitable than rural enterprises and are generally better positioned to attract and cultivate skilled workers. This, in turn, promotes the migration of high-skilled talent to large cities and enhances the exchange of knowledge and skills. In addition, urbanization creates positive externalities for rural areas by facilitating the flow of information, technology, capital, and human resources through migration and remittances. Economic growth and geographic agglomeration are therefore mutually reinforcing, and spatial agglomeration is widely considered conducive to development. In this context, the relationship between financial agglomeration and urbanization has become an important topic for understanding the sustainable development of low-carbon urbanization.^{[9] [10]}

Financial agglomeration contributes to the sustainable development of urbanization through multiple mechanisms, including labor attraction, income growth, capital allocation, information transmission, industrial upgrading, and infrastructure development. These channels not only accelerate the urbanization process but also support coordinated and sustainable urban development in economic, social, and environmental dimensions. Existing research on the impact of financial agglomeration on urbanization has mainly concentrated on several key regional contexts. For example, studies of the Yangtze River Economic Belt have shown that financial agglomeration exerts a smaller positive effect on population urbanization and economic urbanization in the western region than in the eastern and central regions, where the positive effects are more pronounced. Although such studies distinguish among eastern, central, and western China, the geographical scope remains too broad, and the linear distribution of cities along the Yangtze River Economic Belt limits the ability of the findings to fully capture the spatial relationship between regional financial agglomeration and urbanization. To address this limitation, some scholars have focused specifically on eastern China and found that financial agglomeration significantly promotes industrial urbanization and spatial urbanization, whereas its impact on population urbanization is relatively limited. In the Pearl River Delta urban agglomeration, industrial agglomeration has been shown to significantly promote local new urbanization, with notable spatial spillover effects generated by producer services and financial industries. In addition, land urbanization under excessive financial support has been linked to financial efficiency, and excessively high financial efficiency may lead to misallocation and waste of financial resources. At present, research on the relationship between financial agglomeration and urbanization remains concentrated in central and eastern China, where the level of financial agglomeration is relatively high. By contrast, comparatively little attention has been paid to southwestern China.^{[11] [12]}

Against this background, examining the impact of financial agglomeration on urbanization in Southwest China has important theoretical and practical value. First, it supplements the existing literature by shifting attention from developed regions to a less developed area that has received limited scholarly attention, thereby enriching research in financial economics and regional economics. Second, the use of satellite data to measure urbanization levels, combined with denoising techniques, helps improve objectivity and accuracy while reducing potential measurement bias. Third, Southwest China provides a distinctive case for analysis because of its uneven economic development, allowing for a clearer understanding of regional variation under different levels of development. Fourth, although the urbanization rate in Southwest China has risen steadily in recent years, it remains well below the national average. Investigating the spatial relationship between financial agglomeration and urbanization in this region can therefore provide useful evidence for government decision-making, particularly in relation to industrial restructuring and regional development policy.^[13-15]

The remainder of this paper is organized as follows. Section 2 describes the study area and data. Section 3 presents the methodology and related tests. Section 4 reports the regression results. Section 5 provides the robustness analysis. Section 6 concludes the paper and discusses the policy implications. Section 7 outlines the limitations of the study and directions for future research.

2. Research Area and Materials

2.1 Research Area

China's administrative system is characterized by a nested hierarchical structure. Below the central government, the administrative hierarchy descends through provinces, prefecture-level cities, counties, towns, and villages. At the provincial level, China has 34 administrative regions, including 23 provinces, 5 autonomous regions, 4 municipalities, and 2 special administrative regions. Existing studies on financial agglomeration and urbanization have mainly concentrated on the central and eastern regions, as well as several economically advanced belts. In contrast, relatively limited attention has been given to the southwestern region, where financial agglomeration developed comparatively later.

The three southwestern provinces of Sichuan, Guizhou, and Yunnan have a combined population of more than 200 million, accounting for nearly 15 percent of the national population. This large population base provides considerable potential for urbanization. At the same time, the region occupies a strategically important position in major national development initiatives, particularly the Western Development Strategy and the Belt and Road Initiative. Examining the impact of financial agglomeration on urbanization in this region can therefore contribute to improving the quality of regional economic development and strengthening its international competitiveness. As industrial restructuring and upgrading continue, the financial sector has become increasingly important in these provinces and now plays a vital role in promoting regional economic growth. Against this background, this study focuses on 46 prefecture-level cities located in the southwestern provinces of Sichuan, Guizhou, and Yunnan.

2.2 Materials

Satellite data were used to construct indicators of economic urbanization and population urbanization for 46 cities in Southwest China over the period 2011 to 2020. Specifically, nighttime light data were used as a proxy for the level of economic urbanization, whereas urban scale data, following preprocessing and noise reduction procedures, were employed to measure the level of population urbanization.

This study conceptualizes urbanization as comprising two dimensions, namely population urbanization and economic urbanization, which are used as the dependent variables. The core explanatory variable is the level of financial agglomeration, measured by the location entropy index. Originally developed to capture the uneven spatial distribution of economic activity across regions in the United States, the location entropy index has been widely applied in regional economic analysis and urban planning to assess the relative concentration of a particular industry within a specific area.

The financial agglomeration index is calculated as follows:

$$Finsaij = \frac{L_{it}/G_{it}}{L_t/G_t}$$

where L_{it} denotes the total deposits and loans of financial institutions in city i in year t , and L_t denotes the total deposits and loans of all cities in year t . G_{it} represents the GDP of city i in year t , while G_t represents the GDP of all cities in year t . Accordingly, $Finsagg_{it}$ indicates the financial agglomeration index of city i in year t , also referred to as the location entropy of financial agglomeration.

As shown in Table 1, and following previous studies, this paper includes transportation infrastructure, fixed asset investment, foreign direct investment, government expenditure, industrial development, science and technology, education, and healthcare as control variables. Including these control variables helps provide a more comprehensive and accurate examination of the relationship between financial agglomeration and urbanization. By accounting for these important factors, the analysis is better able to identify the independent effect of financial agglomeration on urbanization while also reducing the risk of omitted variable bias.

Table 1 List of variables

	Variable name	Variable symbol	Variable definition	Unit
Explained variables	Population urbanization	Purba	The proportion of urban population to total people	
	Economic urbanization	Eurba	Disposable income of urban residents	

	Variable name	Variable symbol	Variable definition	Unit
Explanatory variables	Financial agglomeration	Finsa	The level of financial agglomeration	/
	Transportation infrastructure	INT	Highway mileage	km
	Fixed assets investment	INF	Fixed assets investment	100 million yuan
	Foreign actual investment	FDI	Actual use of foreign capital	10 thousand dollar
	Government expenditure	GOV	Local general public budget expenditure	10 thousand yuan
Control variables	Industrial development	IND	The second industrial added value	100 million yuan
	Science and technology level	SIC	Internal expenditure on research and development	10 thousand yuan
	Educational level	EDU	Number of beds in Health Institutions	beds
	Medical health level	MED	Number of students enrolled in middle and higher education institutions	people

The study covers 46 prefecture-level cities in Yunnan, Guizhou, and Sichuan provinces from 2011 to 2020. Urbanization levels are derived from satellite data, while economic indicators come from regional and city-level statistical yearbooks, including financial loans, GDP, highway mileage, fixed asset investment, foreign capital usage, R&D expenditure, fiscal spending, secondary industry value added, university enrollment, and hospital beds. For cities with missing data on foreign capital, R&D expenditure, and hospital beds, the moving average smoothing method was used to impute values and ensure data consistency.

3. Methodology

3.1 Moran's I Index

Spatial auto-correlation analysis is a spatial statistical method used to study spatial correlation, reflecting the degree of correlation between certain geographical phenomena or attribute values across regions. Global spatial auto-correlation can indicate the overall characteristics of spatial correlation^{[6][7]}. The global Moran's I index can be expressed as follows:

$$\text{Moran's I} = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{S^2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}}$$

In the formula, $S^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$, $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$, Y_i denotes the financial agglomeration level of city i , n is the total number of cities, \bar{Y} denotes the mean of financial agglomeration levels across all cities. The Moran's I index ranges from -1 to 1: if Moran's $I < 0$, there is negative spatial interaction between regions; if Moran's $I > 0$, there is positive spatial interaction between regions; if Moran's $I = 0$, there is no spatial interaction between regions.

The spatial weight matrix for geographic distance in this study is established based on the inverse square of the geographic distance between two cities, aiming to assess whether geographic proximity influences the impact of financial agglomeration on urbanization. Given data availability, linear distance (d) is used as the standard, as shown below.

$$W_{ij} = \begin{cases} \frac{1}{d_{ij}^2} & (i \neq j) \\ 0 & (i = j) \end{cases}$$

W_{ij} represents the weight of the geographic distance between cities i and j . The smaller the distance between the two cities, the greater the weight. d_{ij} represents the geographic distance between cities i and j .

3.2 Spatial econometric models

Spatial econometric models include the Spatial Lag Model (SLM), Spatial Error Model (SEM), and Spatial Durbin Model (SDM). To determine whether SLM or SEM is more appropriate, the LM (lag) test is proposed. If LM (lag) shows statistical significance in the spatial dependence test, SLM is suitable; if not, SEM is more appropriate. The SLM primarily examines

whether there are spillover effects among variables within regions. The expression is as follows:

$$Y_{it} = \rho \sum_{j=1}^N W_{ij} Y_{it} + \beta X_{it} + \alpha_i + \tau_t + \mu_{it}$$

In this formula, Y_{it} represents the depended variable, referring to the urbanization level of region i in year t ; $\rho \sum_{j=1}^N W_{ij} Y_{it}$ denotes the spatial lag of the dependent variable. X_{it} signifies the core explanatory variable; ρ is the spatial auto-correlation coefficient; W_{ij} is an $n \times n$ matrix representing the spatial weight matrix based on geographic distances, β represents the parameter estimates of the independent variables, α_i stands for spatial fixed effects, τ_t denotes time fixed effects, and μ_{it} represents the random error term.

The spatial error model is suitable for analyzing spatial heterogeneity, where the spatial correlation of the target variable depends on the random error term. In equation, ε_{it} represents the spatial dependence of sample observations, indicating the extent to which variables in one region spillover to neighboring regions. ε_{it} follows a normal distribution as the random error vector. The model can be expressed as:

$$Y_{it} = \beta X_{it} + \alpha_i + \tau_t + \varepsilon_{it}$$

$$\varepsilon_{it} = \lambda \sum_{j=1}^N W_{ij} \varepsilon_{it} + \mu_{it}$$

The Spatial Durbin Model (SDM) is a general form of spatial econometric models. Under certain conditions, it can be simplified to include both spatial lag and spatial error terms of the dependent variable. The model considers not only the spatial correlation of the explanatory variables but also their spatial lag effects. When $\theta = 0$, SDM reduces to the Spatial Lag Model (SLM). when $\theta = \rho = 0$, SDM degrades to the Spatial Error Model (SEM).

The model can be expressed as equation (5):

$$Y_{it} = \rho \sum_{j=1}^N W_{ij} Y_{it} + \beta X_{it} + \theta \sum_{j=1}^N W_{ij} X_{it} + \alpha_i + \tau_t + \mu_{it}$$

In the equation, Y_{it} represents the dependent variable, X_{it} denotes the explanatory variables, ρ is the spatial auto-correlation coefficient, W_{ij} is the spatial weight matrix, $\theta \sum_{j=1}^N W_{ij} X_{it}$ and $\rho \sum_{j=1}^N W_{ij} Y_{it}$ respectively represent the spatial lag of the explanatory and dependent variables. β , θ represent the regression coefficients of the explanatory variables and their spatial relationships, and μ_{it} is the error term. The determination of whether the Spatial Durbin Model (SDM) can be simplified to the Spatial Lag Model (SLM) or Spatial Error Model (SEM) is assessed using Wald tests and LR statistics. If the Wald-spatial-lag and LR-spatial-lag values do not pass significance tests, it indicates that SDM should be reduced to SLM or SEM.

To investigate the spatial relationship between financial agglomeration level and urbanization (both population and economic) in the southwestern region of China, this study introduces a spatial panel Durbin model. All variables are expressed in logarithmic form to mitigate potential heteroskedasticity.

$$\ln Eurba_{it} = \alpha_0 + \rho \sum_{j=1}^n W_{ij} \times (\ln Eurbz_{jt}) + \phi \ln Finsa_{it} + \theta \sum_{j=1}^n W_{ij} \times \ln Finsa_{jt} + \emptyset \ln X_{it} +$$

$$\gamma \sum_{j=1}^n W_{ij} \times \ln X_{jt} + \alpha_i + \tau_t + \mu_{it}$$

$$\ln Purba_{it} = \alpha_0 + \rho \sum_{j=1}^n W_{ij} \times (\ln Purbz_{jt}) + \phi \ln Finsa_{it} + \theta \sum_{j=1}^n W_{ij} \times \ln Finsa_{jt} + \emptyset \ln X_{it} +$$

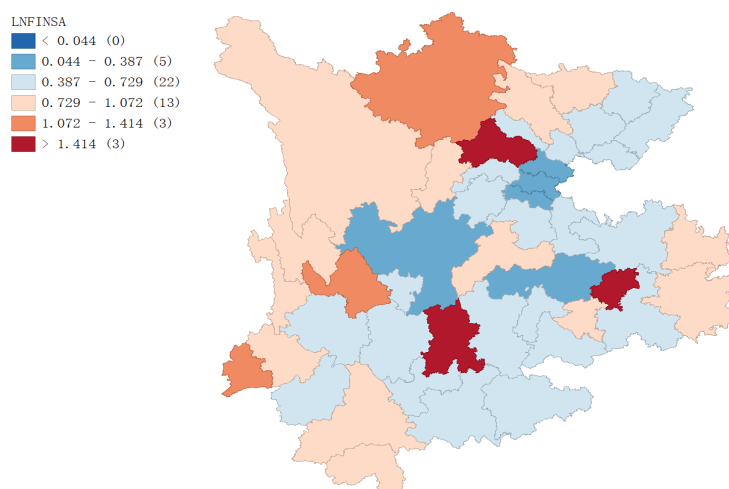
$$\gamma \sum_{j=1}^n W_{ij} \times \ln X_{jt} + \alpha_i + \tau_t + \mu_{it}$$

In the above formula, W represents the spatial matrix where i and j denote different cities; t denotes the year; $Eurba_{it}$, $Purba_{it}$ respectively represent the economic urbanization and population urbanization levels of city i in year t . α_0 is the intercept term; ρ denotes the coefficient of spatial lag term; ϕ , θ and \emptyset , γ are parameters to be estimated. α_i , τ_t respectively represent region fixed effects and time fixed effects; μ_{it} is the error term; $Finsa_{it}$ represents the financial agglomeration level of city i in year t ; X_{it} denotes the set of all control variables.

4. Rebounding and Defensive Comprehensiveness

4.1 Map of Financial Agglomeration Index

Figure 1 The distribution map of the Financial Agglomeration Index for the year 2020.



Based on the data presented in the figure, the distribution of the Financial Agglomeration Index across three provinces in 2020 reveals several key points. First, high financial agglomeration areas in Yunnan and Sichuan are concentrated in provincial capitals and surrounding regions, such as Kunming and Chengdu, with financial agglomeration indices greater than 1.41. In contrast, Guizhou's high financial agglomeration areas are primarily located in and around Guiyang. Secondly, Guizhou's financial agglomeration index distribution is relatively uniform, with most regions exhibiting low levels of financial agglomeration. In comparison, Yunnan and Sichuan display significant variations in their financial agglomeration indices, with distinct high agglomeration zones. Lastly, the peripheral areas of all three provinces generally have lower financial agglomeration indices, indicating that financial resources are predominantly concentrated in central cities and more economically developed regions.

4.2 Spatial Auto-correlation Analysis

Using the inverse distance squared geographic spatial weight matrix and GeoDa software, the bi-variate Moran's I indices for financial agglomeration with population urbanization and economic urbanization in the southwestern region of China from 2011 to 2020 were calculated. Spatial correlation tests were conducted, and the results are shown in Table 2.

Table 2 Moran's I index

Financial agglomeration and population urbanization			Financial agglomeration and economic urbanization		
Year	Moran's I	P-value	Year	Moran's I	P-value
2011	-0.092	0.033	2011	-0.092	0.038
2012	-0.084	0.046	2012	-0.088	0.039
2013	-0.088	0.044	2013	-0.045	0.212
2014	-0.078	0.059	2014	-0.040	0.258
2015	-0.079	0.041	2015	-0.063	0.090
2016	-0.066	0.076	2016	-0.085	0.046
2017	-0.061	0.094	2017	-0.090	0.037
2018	-0.053	0.149	2018	-0.077	0.050
2019	-0.024	0.380	2019	-0.062	0.094
2020	0.048	0.048	2020	0.004	0.376

The results indicate that from 2011 to 2020, the Moran's I statistics for financial agglomeration and urbanization were predominantly negative and statistically significant in most years. This suggests a significant spatial relationship between financial agglomeration and both population urbanization and economic urbanization. Therefore, when analyzing the characteristics and influencing factors of financial agglomeration in the southwestern region of China, it is essential to consider geographic factors and spatial effects. This necessitates the construction of a spatial panel model for further research,

providing a basis for subsequent studies.

4.3 Tests for Spatial Econometric Models

To further examine the presence of spatial correlation and to select the appropriate model, we conducted LM tests, LR tests, and Wald tests in table 3. LM tests are preliminary tests used to determine the existence of spatial auto-correlation. These tests include examinations of the Moran's I statistic as well as tests related to spatial lag and spatial error models. LR tests, on the other hand, are used to assess the suitability of different spatial models, determining which model is the most appropriate. Wald tests are employed to check whether the chosen optimal spatial model could degenerate into other spatial models, ensuring the optimality of the model.

Table 3 Spatial panel model test

Spatial panel model test		Finsa / Eurba		Finsa / Purba	
		Statistic	P-value	Statistic	P-value
LM test	Moran's I	15.298***	0.000	10.138***	0.000
	LM-lag	221.495***	0.000	95.958***	0.000
	Robust-LM-lag	144.499***	0.000	113.981***	0.000
	LM-error	80.278***	0.000	7.575***	0.006
	Roust-LM-error	3.282*	0.070	25.594***	0.000
LR test	LR-SDM/SAR	54.01***	0.000	54.50***	0.000
	LR-SDM/SEM	71.14***	0.000	108.72***	0.000
Wald test	Wald-SDM/SAR	54.87***	0.000	51.99***	0.000
	Wald-SDM/SEM	71.35***	0.000	85.84***	0.000
Spatial-time fixed effect test	LR-both/ind	27.95***	0.0018	3.78	0.9568
	LR-both/time	759.29***	0.0000	2428.58***	0.000

Note: ***, **, * are significant at 1%, 5%, and 10% levels respectively.

The results of the LM test, as shown in the table, further confirm the spatial relationship between financial agglomeration and urbanization. To explore whether the Spatial Durbin Model (SDM) may degenerate into the Spatial Auto-regressive Model (SAR) or the Spatial Error Model (SEM), the results of the LR test indicate that the SDM is more accurate than the SAR and SEM, rejecting the possibility of degeneration into SAR and SEM at the 1% significance level. Additionally, the Wald test results also confirm that the SDM does not degenerate into the SAR or SEM, further validating the robustness of the results.

4.4 Empirical Analysis and Results

Table 4 presents the empirical results and related indicators of the impact of financial agglomeration on population urbanization and economic urbanization under the inverse distance squared geographic spatial weight matrix.

Table 4 The empirical results of SDM

	LnEurba	LnPurba
LnFinsa	0.7063*** (4.52)	0.1448*** (3.30)
LnINT	0.3723*** (4.10)	-0.0294 (-1.16)
LnINF	0.0228 (1.16)	0.0257*** (4.67)
LnFDI	0.0024 (0.71)	-0.0009 (-1.01)
LnGOV	0.1298*** (2.30)	-0.0112 (-0.74)

	LnEurba	LnPurba
LnIND	0.2837*** (5.31)	0.009 (0.60)
LnSIC	0.0091 (0.72)	-0.0059 (-1.64)
LnEDU	0.0292 (1.06)	-0.0208** (-2.69)
LnMED	0.0590 (1.53)	0.0270** (2.50)
WLnFinsa	0.8651*** (2.79)	0.3421*** (4.32)
WLnINT	-0.1067 (-0.49)	0.1450*** (3.71)
WLnINF	0.0200 (0.44)	0.0225* (1.90)
WLnFDI	0.0071 (0.82)	-0.0010 (-0.51)
WLnGOV	0.1163 (1.41)	-0.0283 (-1.36)
WLnIND	-0.4201*** (-3.93)	0.0561** (1.98)
WLnSIC	-0.1058*** (-3.31)	-0.0100 (-1.21)
WLnEDU	0.1338*** (2.67)	-0.0710*** (-4.69)
WLnMED	0.679 (0.77)	0.0360 (1.62)
ρ	0.6082*** (13.59)	0.2369*** (3.61)
R ²	0.7617	0.7283
Log-likelihood	377.91	969.84

Note: ***, **, * are significant at 1%, 5%, and 10% levels respectively.

In this study, the Spatial Durbin Model (SDM) was applied to empirically analyze the relationship between financial agglomeration and urbanization. To enhance the robustness of the estimation results, eight control variables were included: infrastructure level, fixed asset investment, actual foreign investment, government expenditure, industrial level, technological development level, education level, and healthcare level. The results indicate that at the 1% significance level, financial agglomeration positively affects both the internal and spillover effects on population urbanization and economic urbanization. This demonstrates that an increase in the level of financial agglomeration not only promotes population and economic urbanization within the region but also has a positive impact on the urbanization of surrounding areas. In terms of the extent of the impact, the increase in financial agglomeration level has a greater effect on economic urbanization than on population urbanization.

The level of transportation infrastructure shows a positive correlation with the internal effects on economic urbanization. This is primarily because improvements in transportation facilities can effectively promote local economic activities and attract investment. However, its internal effects on population urbanization are insignificantly negative, possibly due to the convenience of transportation leading people to choose to live in surrounding areas. Fixed asset investment shows a positive correlation with both population urbanization and economic urbanization, but its positive correlation with economic

urbanization is not significant. This is mainly due to weak economic foundations, industrial structure issues, and low capital efficiency. Actual foreign investment shows an insignificantly negative correlation with both economic and population urbanization. This reflects the limited effect of foreign investment in driving economic restructuring and development. Government expenditure shows a significantly positive correlation with the internal effects on economic urbanization, mainly attributed to its positive role in infrastructure development and job creation. However, it shows an insignificantly negative correlation with the internal effects on population urbanization, indicating that government spending has not effectively promoted population migration and integration into urban areas. The development of the secondary industry has a significantly positive correlation with the internal effects on economic urbanization, while it shows an insignificantly positive correlation with population urbanization. This may be due to the diversity of the labor source and the mobility of local residents. Research and development (R&D) expenditure shows an insignificantly positive correlation with the internal effects on economic urbanization and an insignificantly negative correlation with population urbanization. This is primarily because the effects of technological innovation take time to materialize and are influenced by the mobility of talent. Education level shows an insignificantly positive correlation with the internal effects on economic urbanization, possibly due to the mismatch in the job market and economic structural limitations. However, it shows a significantly negative correlation with the internal effects on population urbanization, mainly due to the outflow of highly educated individuals. Healthcare level shows an insignificantly positive correlation with the internal effects on economic urbanization, but a significantly positive correlation with the internal effects on population urbanization. This reflects the attractiveness of high-level medical services to residents' quality of life and settlement choices.

The level of transportation infrastructure shows an insignificantly negative correlation with the spillover effects on economic urbanization in surrounding areas, possibly due to resource competition and regional development imbalances. However, it has a significantly positive correlation with the spillover effects on population urbanization, reflecting the positive influence of improved transportation facilities on population inflow and diverse residential choices in neighboring regions. Fixed asset investment shows a positive correlation with both economic and population urbanization spillover effects. However, the positive correlation with economic urbanization is not significant, mainly due to weak economic foundations and industrial structure issues. The positive correlation with population urbanization is significant, benefiting from improvements in infrastructure and public services, as well as the promotion of regional coordinated development. Actual foreign investment shows insignificant spillover effects on both economic and population urbanization. This indicates the limited impact of foreign investment in balancing regional development and increasing the inflow of external populations. Government expenditure shows a positive but insignificant spillover effect on economic urbanization, possibly influenced by resource constraints and policy implementation issues. The spillover effect on population urbanization is insignificantly negative, suggesting that government expenditure might increase the pressure on urban social management. The development of the secondary industry has a significantly negative spillover effect on economic urbanization, mainly due to environmental burdens and resource competition. However, it has a significantly positive spillover effect on population urbanization, primarily due to the increase in social service facilities and the driving force of regional economic growth. R&D expenditure shows a significantly negative spillover effect on economic urbanization, reflecting the impact of technological diffusion on local economies. The spillover effect on population urbanization is insignificantly negative, mainly due to the mobility of talent. Education level shows a significantly positive spillover effect on economic urbanization, highlighting its positive role in attracting talent and promoting technology transfer. However, it has a significantly negative spillover effect on population urbanization, possibly due to the outflow and settlement trends of highly educated individuals. Healthcare level shows an insignificantly positive spillover effect on economic urbanization and an insignificantly positive spillover effect on population urbanization. This may be because the limitations of healthcare services have not significantly extended to surrounding areas.

5. Conclusions and Recommendations

This study focuses on 46 prefecture-level cities in southwestern China from 2011 to 2020, using the Location Quotient (LQ) method to measure the level of financial agglomeration and employing satellite data to estimate the levels of population and economic urbanization. Moran's I index was used for spatial auto-correlation analysis, followed by the Spatial Durbin Model

(SDM) to investigate the spatial relationship between financial agglomeration and urbanization. Transportation infrastructure, healthcare levels, fixed asset investment, industrial development, technological advancement, and education levels were included as control variables in the empirical analysis. The results indicate that financial agglomeration has significant positive internal and spillover effects on population and economic urbanization at the 1% significance level. The main conclusions and policy recommendations are as follows.

5.1 Main Conclusions

Spatial Characteristics of Financial Agglomeration: Southwestern China exhibits spatial characteristics of financial agglomeration. Financial agglomeration has significant positive internal and spillover effects on both population and economic urbanization at the 1% significance level, primarily influencing urbanization levels through indirect effects.

Control Variables' Influence: Industrial development, government expenditure, fixed asset investment, and healthcare levels significantly promote both economic and population urbanization in southwestern China. Education levels negatively affect the internal effects on population urbanization but positively influence the spillover effects on economic urbanization. Transportation infrastructure and healthcare levels mainly impact urbanization through direct effects, while industrial development, government expenditure, fixed asset investment, education levels, and technological advancement mainly affect urbanization through indirect effects.

5.2 Policy Recommendations

Promote Regional Financial Centers: Develop regional financial centers in southwestern China by selecting cities with geographical and economic advantages to establish financial special zones, attracting domestic and international financial institutions. Enhance financial infrastructure, such as financial technology, payment systems, and credit systems, in less developed areas to improve financial service convenience and efficiency. Provide tax incentives, financing facilitation, and land use benefits to attract financial institutions and related enterprises.

Strengthen Financial Support for Urbanization: Innovate financial products and services to meet urbanization needs, such as housing loans and urban infrastructure financing, providing financial support for both population and economic urbanization. Promote inclusive finance by extending financial services to small and medium-sized cities and towns through policy guidance and technological innovation to ensure balanced regional development.

(1) **Leverage Financial Agglomeration Spillover Effects:** Enhance regional financial cooperation to facilitate cross-regional financial resource flows, thereby boosting the spillover effects of financial agglomeration. Establish regional financial cooperation mechanisms for resource sharing and complementary advantages. Increase investment in financial technology R&D, encourage the use of big data and AI by financial institutions to improve service efficiency, and focus on financial talent development to attract high-end financial professionals and elevate the regional financial sector.

(2) **Invest in Infrastructure and Public Services:** Increase investments in transportation, healthcare, government spending, industrial development, fixed assets, and education to provide a solid foundation for urbanization through their internal and spillover effects. Pay special attention to underdeveloped areas in southwestern China, utilizing these elements to maximize their spillover effects on surrounding regions.

(3) **Enhance Policy Coordination and Innovation:** Establish multi-level coordination mechanisms involving local governments, financial institutions, and enterprises to form a cohesive policy force for promoting financial agglomeration and urbanization. Deepen financial system reforms, open up financial markets, and improve the competitiveness and service level of the financial sector to support the urbanization process. Continuously monitor and evaluate the development of urbanization and financial agglomeration, adjusting policies as needed to ensure effective implementation and achievement of objectives.

(4) **Implementing these policy recommendations** can effectively promote the coordinated development of financial agglomeration, population urbanization, and economic urbanization in southwestern China, thereby enhancing the region's overall competitiveness and sustainability.

Funding

No

Conflict of Interests

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

Reference

- [1] Xuan, Z., Guanghe, R., & Ke, C. (2020). A study of the spatial effects of financial agglomeration on regional economic growth of private sectors. *Science Research Management*, 41(5), 259.
- [2] Krugman, P. (1991). Increasing returns and economic geography. *Journal of political economy*, 99(3), 483-499.
- [3] Porter, M. E. (1990). New global strategies for competitive advantage. *Planning review*, 18(3), 4-14.
- [4] Wang, X., Zhu, L., Li, N., Wang, S., & Yang, S. (2023). A study of the impact of financial agglomeration on the urban-rural income gap. *Financial Engineering and Risk Management*, 6(11), 92-101.
- [5] Wen, Y., Zhao, M., Zheng, L., Yang, Y., & Yang, X. (2023). Impacts of financial agglomeration on technological innovation: a spatial and nonlinear perspective. *Technology Analysis & Strategic Management*, 35(1), 17-29.
- [6] Qin, D., & Puttanapong, N. (2025). Estimating county-level digital economy development in china using satellite data: insights from explainable spatial machine learning. *Geocarto International*, 40(1), 2596971. <https://doi.org/10.1080/10106049.2025.2596971>
- [7] Qin, D., & Puttanapong, N. (2026). Evaluating LJ1-01 nighttime light data for economic analysis at China's county level. *Remote Sensing Letters*, 17(3), 324-335. <https://doi.org/10.1080/2150704X.2026.2629349>
- [8] Hu, Y., Wang, C., Zhang, X., Wan, H., & Jiang, D. (2023). Financial agglomeration and regional green innovation efficiency from the perspective of spatial spillover. *Journal of Innovation & Knowledge*, 8(4), 100434. <https://doi.org/10.1016/j.jik.2023.100434>
- [9] Gabriel, S. A., & Rosenthal, S. S. (2013). Urbanization, agglomeration economies, and access to mortgage credit. *Regional Science and Urban Economics*, 43(1), 42-50. <https://doi.org/10.1016/j.regsciurbeco.2012.11.006>
- [10] Che, X., Li, B., Guo, K., & Wang, J. (2011). Property prices and bank lending: some evidence from China's Regional Financial Centres. *Procedia Computer Science*, 4, 1660-1667.
- [11] Ruiz, M., Martínez, J. A., López-Hernández, F., & Castellano, A. (2014). The relationship between concentration of scoring and offensive efficiency in the NBA. *Journal of Quantitative Analysis in Sports*, 10(1), 27-36. <https://doi.org/10.1515/jqas-2013-0062>
- [12] Feng, Y., Zou, L., Yuan, H., & Dai, L. (2022). The spatial spillover effects and impact paths of financial agglomeration on green development: Evidence from 285 prefecture-level cities in China. *Journal of cleaner production*, 340, 130816.
- [13] Yuan, H., Feng, Y., Lee, J., Liu, H., & Li, R. (2020). The spatial threshold effect and its regional boundary of financial agglomeration on green development: a case study in China. *Journal of Cleaner Production*, 244, 118670.
- [14] Levine, R. (1997). Financial development and economic growth: views and agenda. *Journal of economic literature*, 35(2), 688-726.
- [15] Turok, I., & McGranahan, G. (2013). Urbanization and economic growth: the arguments and evidence for Africa and Asia. *Environment and urbanization*, 25(2), 465-482.

Value-Building Rather Than Risk-Reduction? Effects of TikTok Beauty Recommendation Videos on College Students' Purchase Decisions

Qiuyan Li, Fengyu Zhao*

College of Economics and Management, Southwest Forestry University, Kunming, Yunnan, 650224, China

*Corresponding author: Fengyu Zhao, zhaofengyu@swfu.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: TikTok-style short-form video platforms have become an important channel for beauty product discovery and social commerce among young consumers. However, existing studies have mainly examined general video attributes such as informativeness, entertainment, and trust, while paying less attention to beauty-specific content characteristics and the comparative roles of perceived value and perceived risk. This study examines how three content characteristics of TikTok beauty recommendation videos - professionalism, authenticity, and practicality - influence college students' purchase decisions. Drawing on the stimulus-organism-response framework, the study argues that these content characteristics affect purchase decisions through perceived value and perceived risk, while consumption values may further shape these relationships. A questionnaire survey was conducted among Chinese college students who had watched beauty recommendation videos on Douyin, the Chinese counterpart of TikTok, and the data were analyzed using reliability and validity tests, correlation analysis, regression analysis, mediation analysis, and exploratory moderation analysis. The findings show that professionalism, authenticity, and practicality all positively influence purchase decisions. Perceived value plays a significant mediating role, whereas perceived risk does not show a significant mediating effect. Additional analysis indicates that hedonic consumption values strengthen the effect of professionalism on perceived value but weaken the effect of authenticity on perceived value. The study contributes to short-form video marketing research by refining beauty-related content characteristics and by showing that value-building is a more important mechanism than risk-reduction in this social commerce context.

Keywords: TikTok; Douyin; Beauty Recommendation Videos; Content Characteristics; Perceived Value; Perceived Risk; Purchase Decision

Published: Apr 24, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1310>

1. Introduction

Social media marketing has evolved from static brand communication to highly interactive, creator-centered, and video-based persuasion. Among emerging formats, short-form video has become especially influential because it combines visual demonstration, emotional appeal, algorithmic distribution, and social interaction in a single environment. In beauty consumption, such features are particularly important. Consumers often rely on creators' demonstrations, product comparisons, and usage experiences to evaluate whether a product is worth buying. As a result, TikTok and its Chinese

counterpart Douyin have become major channels through which young consumers discover, evaluate, and purchase beauty products^[1].

Prior research has shown that influencer marketing can significantly shape consumer attitudes, trust formation, and purchase-related intentions. Lou and Yuan found that informative value and source credibility in influencer content enhance trust in branded social media posts, which then improves brand awareness and purchase intention^[1]. Liu and Zheng further showed that informativeness, authenticity, and homophily strengthen parasocial relationships, which in turn increase brand credibility and purchase intention^[2]. In the beauty domain, influencer credibility appears especially important. Garg and Bakshi reported that beauty vloggers' trustworthiness, expertise, and attractiveness build follower trust and increase purchase intention, with trust acting as a partial mediator^[3]. These findings suggest that beauty-related social media persuasion depends not only on exposure, but also on how consumers interpret the quality and credibility of content.

Recent studies on short-form video marketing have extended this line of inquiry to platform-based social commerce contexts. Luo et al. demonstrated that content features such as usefulness, ease of use, and entertainment value increase consumer trust and perceived value on short-video platforms, which then promote purchase intention^[4]. Shen and Wang showed that users' perceptions of creators' personas affect purchase intention through shared value creation in short-video social commerce^[5]. Zheng found that flow experience significantly increases participative and sharing behavior on Douyin, indicating that immersive engagement is an important psychological mechanism in short-video use^[6]. A recent review confirmed that research on social media influencers and purchase intention is growing rapidly, but also highlighted the need for more nuanced work on emerging content formats, psychological mechanisms, and contextual differences^[7].

Although the existing literature has generated important insights, at least three limitations remain. First, most studies still rely on relatively general dimensions such as informativeness, entertainment, interaction, or trust. These variables are valuable, but they may not fully capture the logic of beauty recommendation content. In beauty short videos, consumers often pay attention to whether the creator appears professionally knowledgeable, whether the demonstration feels authentic, and whether the content is practically useful for real purchase and usage decisions. These three dimensions - professionalism, authenticity, and practicality - are highly salient in beauty consumption, yet they have received less systematic attention in the short-form video literature.

Second, prior studies have more often emphasized positive mechanisms such as trust, parasocial interaction, perceived value, or shared values. By contrast, the role of perceived risk has received less integrated treatment in this context. This is a notable gap because beauty consumption is inherently uncertain. Consumers may worry about product performance, skin compatibility, ingredient safety, or wasted money. At the same time, beauty recommendation videos may reduce such uncertainty by offering demonstrations, reviews, and comparisons. It is therefore theoretically meaningful to examine perceived value and perceived risk together rather than treating only one psychological pathway as central.

Third, there is still limited evidence focusing specifically on college students in the Chinese Douyin/TikTok beauty context. College students are heavy users of short-form video platforms and active consumers of beauty recommendation content. Yet they also differ from other consumer groups in at least two ways. They are often highly responsive to digital content and social influence, but their purchase decisions are constrained by limited budgets and still-developing consumption experience. This makes them an important group for understanding how short-form beauty content translates into actual decision-making. To address these gaps, this study examines how professionalism, authenticity, and practicality in beauty recommendation videos influence college students' purchase decisions. In this manuscript, TikTok is used as a general label for the short-form video commerce context, whereas the empirical data were collected from Chinese college students in the Douyin environment. Drawing on the stimulus-organism-response framework, the study conceptualizes the three content characteristics as stimuli, perceived value and perceived risk as organism variables, and purchase decision as the behavioral response. It also explores whether utilitarian and hedonic consumption values shape selected relationships. This study makes three main contributions. First, it refines short-form video content research by focusing on beauty-specific content characteristics rather than only generic platform features. Second, it extends the application of the stimulus-organism-response framework by comparing a value-building mechanism with a risk-reduction mechanism in the same model. Third, it

provides evidence from the Douyin/TikTok beauty recommendation context among Chinese college students, thereby adding a more specific social commerce setting to the broader literature on influencer marketing and purchase intention.

Based on the above arguments, this study advances three sets of propositions. First, professionalism, authenticity, and practicality are expected to positively affect purchase decision. Second, these relationships are expected to be transmitted primarily through perceived value and, to a lesser extent, through perceived risk. Third, utilitarian and hedonic consumption values are explored as potential boundary conditions that may shape selected psychological paths.

2. Materials and Methods

2.1 Research Design

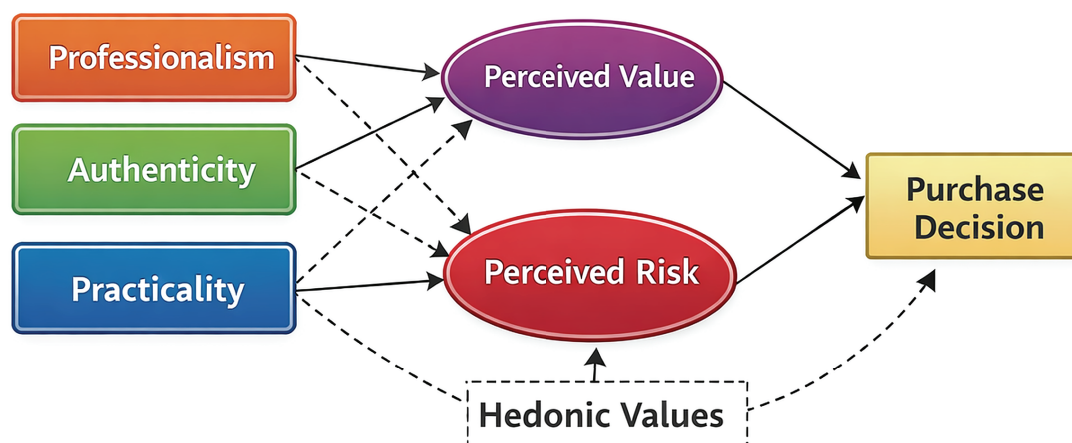
This study employed a quantitative, cross-sectional survey design to examine how the content characteristics of beauty recommendation videos influence college students' purchase decisions. The research model was grounded in the stimulus-organism-response framework, in which professionalism, authenticity, and practicality were treated as stimulus variables; perceived value and perceived risk were modeled as organism variables; and purchase decision was treated as the behavioral response^[8]. In addition, utilitarian and hedonic consumption values were explored as boundary conditions that may shape selected relationships in the model.

A survey design was considered appropriate because the study tested a set of theory-driven relationships among latent variables and focused on respondents' current perceptions and decision tendencies in a specific short-form video commerce context. The questionnaire was developed in two stages. A pilot survey was conducted first to refine item wording and eliminate problematic indicators. The formal survey was then administered to the target population using the revised instrument.

The overall research design is presented in Figure 1.

2.2 Conceptual Framework

Figure 1. Conceptual framework of the study



The conceptual framework places professionalism, authenticity, and practicality on the stimulus side of the model. Perceived value and perceived risk represent organism variables, and purchase decision represents the response variable. Utilitarian and hedonic consumption values are considered boundary conditions in the exploratory moderation analysis.

2.3 Sampling and Data Collection

The target population of this study consisted of Chinese college students who had been exposed to beauty recommendation videos on Douyin. A non-probability convenience sampling strategy was adopted because the study focused on a specific user group and required efficient access to respondents through online channels. The questionnaire was distributed through Tencent Questionnaire, WeChat contacts, WeChat groups, and campus-related online communities. These channels were appropriate because college students are highly active in mobile social media environments and can be reached efficiently through platform-based recruitment. The same online channels were also used during the pilot stage of the study.

Before the formal survey, a pilot test was conducted to improve the content validity and clarity of the instrument. A total of

110 pilot questionnaires were collected, of which 77 were valid and used for preliminary reliability and exploratory factor analyses. Based on the pilot results, several items with low factor loadings, cross-loadings, or weak conceptual fit were removed. This process led to a more concise and stable measurement instrument for the main survey.

The formal survey yielded 372 valid responses for empirical analysis. Respondents were screened to ensure that they were current college students and had experience watching beauty recommendation videos on Douyin/TikTok. The final sample size was adequate for multivariate analysis and mediation testing, given the number of retained observed variables in the measurement model. All items were measured on a five-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Higher scores indicated stronger agreement with each statement and stronger perceived presence of the underlying construct.

2.4 Measures

The questionnaire included eight constructs: professionalism, authenticity, practicality, perceived value, perceived risk, utilitarian consumption values, hedonic consumption values, and purchase decision. The measurement design drew on established studies in influencer marketing, short-form video marketing, consumer value, and consumer risk research^[9]. Item wording was adapted to the beauty recommendation context and then refined through the pilot test.

Professionalism refers to the extent to which creators provide knowledgeable, credible, and product-relevant explanations and recommendations, such as ingredient interpretation, skin-type matching, and usage guidance. Authenticity refers to the extent to which the video presents product effects, usage experiences, and claims in a realistic and non-exaggerated way. Practicality refers to the extent to which the content provides usable and decision-relevant information, such as product comparisons, usage instructions, and actionable purchase suggestions.

Perceived value was conceptualized as the consumer's overall evaluation of whether the recommended product is worth buying after weighing expected benefits against expected costs^{[9][10]}. Perceived risk captured concerns about uncertainty and possible negative outcomes in the purchase decision, especially in relation to product safety, product performance, and financial loss^[11]. Utilitarian consumption values and hedonic consumption values were measured to reflect two different consumption orientations^[12]. Finally, purchase decision tendency was measured through respondents' willingness to purchase products recommended in beauty recommendation videos.

2.5 Instrument Refinement and Data Analysis

As shown in Table 1, the final instrument retained only those items that were conceptually clear and statistically acceptable after the pilot test. Reliability analysis was first conducted for each construct in the pilot stage. An exploratory factor analysis was then conducted using principal component extraction and varimax rotation. Items with low factor loadings, cross-loadings, or weak conceptual alignment with their intended construct were considered for deletion. One item from purchase decision, one item from professionalism, one item from utilitarian consumption values, one item from perceived value, and two items from practicality were removed in this process.

Data analysis was conducted in several steps. First, descriptive statistics were used to summarize the sample characteristics. Second, reliability and validity tests were performed to evaluate the measurement model, including internal consistency, convergent validity, and factor structure. Third, common method bias was assessed because the data were collected through a self-reported questionnaire at a single point in time. Fourth, correlation analysis was used to examine the relationships among the main constructs. Finally, regression-based analyses were conducted to assess direct effects, mediation effects, and exploratory moderation effects. The analyses were performed using SPSS 26.0 and SPSSAU.

Table 1. Summary of constructs and measurement design

Construct	Definition	Items	Example Item
Professionalism	The extent to which the creator provides expert, credible, and product-relevant information	3	The creator can provide professional and objective advice for different skin types and makeup needs.
Authenticity	The extent to which the video presents product effects and experiences in a realistic and non-exaggerated way	4	The video does not use excessive filters or effects to exaggerate the product outcome.

Construct	Definition	Items	Example Item
Practicality	The extent to which the content offers useful and applicable information for real purchase and usage decisions	2	The product comparison or matching advice in the video is useful for me.
Perceived Value	The respondent's overall judgment that the recommended beauty product is worth buying	4	Overall, the product recommended in the video is worth purchasing.
Perceived Risk	The respondent's concern about possible negative outcomes of purchasing the recommended product	6	I worry that this product may cause skin problems, such as allergies.
Utilitarian Consumption Values	A functional and task-oriented tendency in beauty consumption	2	I pay more attention to whether the product can effectively solve a practical problem.
Hedonic Consumption Values	An enjoyment-oriented tendency that emphasizes pleasure and emotional experience in beauty consumption	3	I enjoy the pleasure of trying new beauty products.
Purchase Decision	The respondent's tendency to purchase products recommended in TikTok beauty videos	2	I am willing to buy the beauty product recommended in the video.

3. Results

3.1 Sample Profile

A total of 372 valid responses were included in the final analysis. As shown in Table 2, female respondents accounted for the majority of the sample (83.06%), which is broadly consistent with the audience profile of beauty recommendation content on short-form video platforms. In terms of academic year, sophomores accounted for the largest proportion (44.09%), followed by freshmen (22.04%), juniors (19.62%), and seniors (14.25%). Regarding monthly disposable living expenses, most respondents reported budgets between RMB 1001 and 1500 (41.94%) or between RMB 1501 and 2000 (31.18%). In addition, exposure to beauty recommendation videos was high: 20.43% reported watching such content every day and 50.00% reported watching it frequently. The sample also showed strong behavioral relevance to the research context, as 97.58% of respondents reported that they had engaged in beauty consumption triggered by such videos at least occasionally.

As Table 2 indicates, the respondents were not only frequent short-video users but also highly engaged with beauty recommendation content. This supports the contextual suitability of the sample for testing the proposed model.

Table 2. Sample profile of respondents (N = 372)

Variable	Category	Frequency	Percentage (%)
Gender	Male	63	16.94
	Female	309	83.06
Academic year	Freshman	82	22.04
	Sophomore	164	44.09
	Junior	73	19.62
	Senior	53	14.25
Monthly disposable living expenses	RMB 1000 or below	40	10.75
	RMB 1001-1500	156	41.94
	RMB 1501-2000	116	31.18
	RMB 2001 and above	60	16.13

Variable	Category	Frequency	Percentage (%)
Frequency of watching beauty recommendation videos	Every day	76	20.43
	Frequently	186	50.00
	Occasionally	104	27.96
	Rarely	6	1.61
Beauty consumption triggered by recommendation videos	Many times	172	46.24
	Occasionally	191	51.34
	Never	9	2.42

3.2 Reliability and Validity

The internal consistency of the scales was first assessed using Cronbach's alpha coefficients and corrected item-total correlations. As shown in Table 3, the alpha coefficients ranged from 0.602 to 0.921. Perceived risk (alpha = 0.921), authenticity (alpha = 0.854), and perceived value (alpha = 0.822) showed strong internal consistency. Professionalism (alpha = 0.726), hedonic consumption values (alpha = 0.784), and purchase decision (alpha = 0.726) also reached acceptable levels. Practicality (alpha = 0.651) and utilitarian consumption values (alpha = 0.602) were slightly lower, but both were retained because each scale contained only two items and the retained indicators remained theoretically meaningful and statistically usable.

Construct validity was then examined through KMO and Bartlett's test, as well as confirmatory factor analysis. The KMO value was 0.925 and Bartlett's test of sphericity was significant (chi-square = 5386.878, df = 325, $p < 0.001$), indicating that the dataset was suitable for factor analysis. Table 3 also reports the AVE and CR values. Authenticity, perceived value, perceived risk, hedonic consumption values, and purchase decision all exceeded the conventional thresholds for both CR and AVE. Professionalism showed a CR above 0.7 and an AVE close to 0.5. Practicality and utilitarian consumption values were somewhat lower, but both remained within an acceptable range when considering the small number of retained items and their theoretical relevance. Overall, the measurement model showed acceptable convergent validity.

Table 3. Reliability and convergent validity results

Construct	No. of Items	Cronbach's alpha	AVE	CR
Professionalism	3	0.726	0.472	0.728
Authenticity	4	0.854	0.596	0.855
Practicality	2	0.651	0.482	0.651
Perceived Value	4	0.822	0.540	0.824
Perceived Risk	6	0.921	0.662	0.921
Utilitarian Consumption Values	2	0.602	0.435	0.605
Hedonic Consumption Values	3	0.784	0.550	0.785
Purchase Decision	2	0.726	0.573	0.728

The overall measurement model also demonstrated acceptable fit. As shown in Table 4, the chi-square to degrees-of-freedom ratio was 2.187, RMSEA was 0.057, RMR was 0.041, SRMR was 0.042, and the incremental fit indices CFI, TLI, and IFI were all above 0.9. Although GFI and NFI were slightly below the conventional 0.90 threshold, the overall pattern of fit indices indicates an acceptable model fit.

Because all variables were measured through self-reports in a single questionnaire, common method bias was also tested using Harman's single-factor test. The first unrotated factor explained 34.13% of the total variance, which was below the 40% threshold. This result suggests that common method bias was unlikely to threaten the validity of the findings.

Table 4. Measurement model fit indices

Index	Value
chi-square/df	2.187
GFI	0.888
RMSEA	0.057
RMR	0.041
CFI	0.938
NFI	0.893
TLI	0.926
IFI	0.939
AGFI	0.855
PGFI	0.686
PNFI	0.745
PCFI	0.782
SRMR	0.042

3.3 Descriptive Statistics and Correlations

Descriptive statistics and Pearson correlations are reported in Table 5. Purchase decision had a mean score of 3.949 (SD = 0.739), indicating a moderately positive tendency toward purchasing products recommended in beauty recommendation videos. Among the independent variables, practicality had the highest mean (M = 3.922, SD = 0.725), followed by professionalism (M = 3.804, SD = 0.700) and authenticity (M = 3.538, SD = 0.838).

The correlations were generally consistent with the theoretical expectations. Purchase decision was positively correlated with professionalism ($r = 0.536$, $p < 0.01$), authenticity ($r = 0.543$, $p < 0.01$), practicality ($r = 0.516$, $p < 0.01$), and perceived value ($r = 0.655$, $p < 0.01$). By contrast, perceived risk was not significantly correlated with purchase decision, suggesting that risk perception may be less central than value perception in this context. In addition, professionalism, authenticity, and practicality were all positively correlated with perceived value, while authenticity was negatively correlated with perceived risk ($r = -0.211$, $p < 0.01$). These initial findings provided support for the subsequent regression and mediation analyses.

Table 5. Descriptive statistics and correlations

Variable	Mean	SD	1	2	3	4	5	6	7	8
1. Purchase Decision	3.949	0.739	1							
2. Hedonic Consumption Values	3.744	0.837	0.587**	1						
3. Utilitarian Consumption Values	4.230	0.697	0.342**	0.234**	1					
4. Perceived Risk	3.491	1.046	-0.076	-0.129*	0.226**	1				
5. Perceived Value	3.745	0.755	0.655**	0.542**	0.358**	-0.181**	1			
6. Practicality	3.922	0.725	0.516**	0.446**	0.433**	-0.059	0.612**	1		
7. Authenticity	3.538	0.838	0.543**	0.490**	0.308**	-0.211**	0.742**	0.551**	1	
8. Professionalism	3.804	0.700	0.536**	0.453**	0.387**	-0.094	0.704**	0.637**	0.702**	1

Note. * $p < 0.05$, ** $p < 0.01$.

3.4 Hypothesis Testing

The direct effects of professionalism, authenticity, and practicality on purchase decision were tested using multiple regression analysis. As shown in Table 6, all three content characteristics had significant positive effects on purchase decision. Practicality showed the largest unstandardized coefficient in the direct-effect model ($B = 0.249$, $p < 0.001$), followed closely by authenticity ($B = 0.245$, $p < 0.001$), while professionalism also remained significant ($B = 0.196$, $p = 0.004$). The model explained 37.6% of the variance in purchase decision ($R\text{-squared} = 0.376$). Multicollinearity was not a concern, as all VIF values were below 5.

These findings indicate that higher-quality beauty recommendation content is associated with stronger purchase decision tendencies among college students.

Table 6. Regression results for the direct effects of content characteristics on purchase decision

Predictor	B	SE	Beta	t	p
Constant	1.361	0.185	-	7.366	<0.001
Practicality	0.249	0.055	0.244	4.494	<0.001
Authenticity	0.245	0.052	0.278	4.716	<0.001
Professionalism	0.196	0.067	0.186	2.915	0.004

The next set of analyses examined whether the three content characteristics influenced perceived value and perceived risk. As shown in Table 7, professionalism ($B = 0.283$, $p < 0.001$), authenticity ($B = 0.404$, $p < 0.001$), and practicality ($B = 0.206$, $p < 0.001$) all positively affected perceived value. The explanatory power of the model was relatively high ($R\text{-squared} = 0.639$), suggesting that the three content dimensions are strong predictors of value perception in this context.

Table 7. Regression results for the effects of content characteristics on perceived value

Predictor	B	SE	Beta	t	p
Constant	0.428	0.144	-	2.981	0.003
Professionalism	0.283	0.052	0.262	5.411	<0.001
Authenticity	0.404	0.040	0.449	10.022	<0.001
Practicality	0.206	0.043	0.198	4.791	<0.001

By contrast, the regression on perceived risk in Table 8 shows a more selective pattern. Only authenticity had a significant negative effect on perceived risk ($B = -0.372$, $p < 0.001$), while professionalism and practicality were not significant predictors. The model explained only 5.2% of the variance in perceived risk, indicating that the three content characteristics were much more effective in shaping perceived value than in reducing perceived risk.

Table 8. Regression results for the effects of content characteristics on perceived risk

Predictor	B	SE	Beta	t	p
Constant	4.041	0.322	-	12.532	<0.001
Professionalism	0.123	0.117	0.082	1.049	0.295
Authenticity	-0.372	0.091	-0.298	-4.111	<0.001
Practicality	0.076	0.097	0.053	0.788	0.431

The effects of perceived value and perceived risk on purchase decision were then tested. As shown in Table 9, perceived value had a significant positive effect on purchase decision ($B = 0.648$, $p < 0.001$), whereas perceived risk was not significant ($B = 0.031$, $p = 0.270$).

Table 9. Regression results for the effects of perceived value and perceived risk on purchase decision

Predictor	B	SE	Beta	t	p
Constant	1.413	0.193	-	7.326	<0.001
Perceived Value	0.648	0.039	0.663	16.585	<0.001
Perceived Risk	0.031	0.028	0.044	1.105	0.270

Bootstrap mediation analysis was conducted using 5,000 resamples. The results showed that perceived value significantly mediated the effects of professionalism, authenticity, and practicality on purchase decision, whereas perceived risk did not mediate any of the three relationships. As shown in Table 10, the indirect effects via perceived value were significant because the bootstrap confidence intervals did not include zero.

Table 10. Bootstrap mediation results

Path	Indirect Effect	95% CI	p	Interpretation
Professionalism - Perceived Value - Purchase Decision	0.131	0.073 ~ 0.199	<0.001	Full mediation
Authenticity - Perceived Value - Purchase Decision	0.186	0.124 ~ 0.260	<0.001	Full mediation
Practicality - Perceived Value - Purchase Decision	0.095	0.048 ~ 0.151	<0.001	Partial mediation
Professionalism - Perceived Risk - Purchase Decision	0.004	-0.005 ~ 0.016	0.445	Not supported
Authenticity - Perceived Risk - Purchase Decision	-0.011	-0.033 ~ 0.008	0.225	Not supported
Practicality - Perceived Risk - Purchase Decision	0.002	-0.005 ~ 0.013	0.573	Not supported

More specifically, perceived value fully mediated the effects of professionalism and authenticity on purchase decision. For authenticity, the indirect effect through perceived value was 0.131 (95% CI = 0.073 to 0.199, $p < 0.001$), and the corresponding direct effect became non-significant after perceived value was entered into the model. Practicality, by contrast, retained both a significant indirect effect via perceived value (0.095, 95% CI = 0.048 to 0.151, $p < 0.001$) and a significant direct effect, indicating partial mediation. All indirect paths through perceived risk were non-significant.

Overall, the mediation results indicate that perceived value is the core psychological pathway through which beauty recommendation video content influences purchase decision in this context.

The originally specified moderation assumptions were largely unsupported, especially for utilitarian consumption values and for perceived risk. As shown in Table 11, none of the interaction terms between content characteristics and either utilitarian or hedonic consumption values significantly predicted perceived risk. This suggests that the relationship between content characteristics and perceived risk was relatively stable across different consumption value orientations.

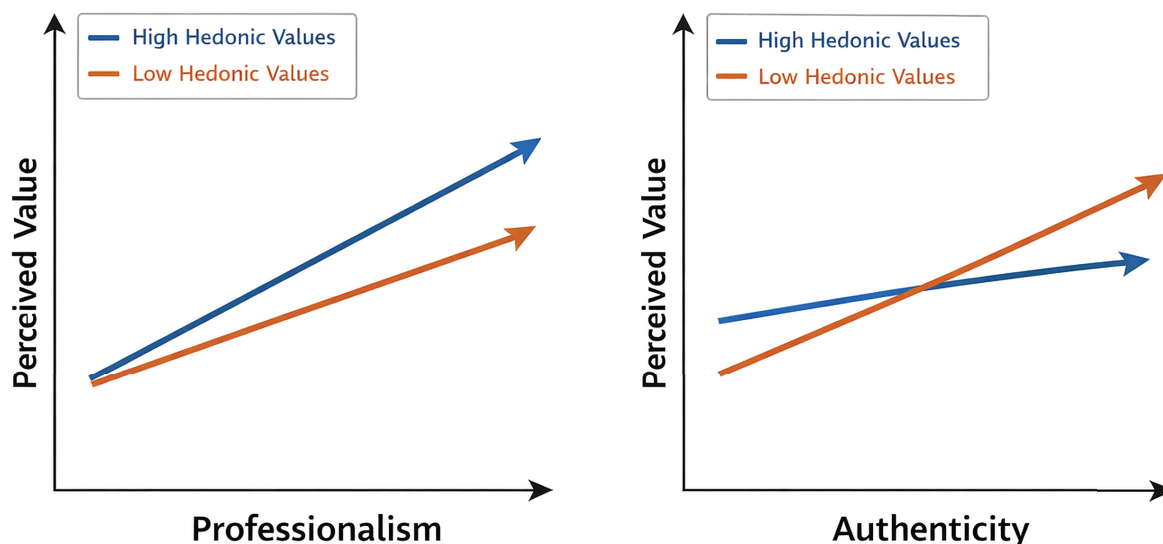
However, additional interaction analysis for perceived value yielded two significant findings. First, the interaction between professionalism and hedonic consumption values was positive and significant ($B = 0.126$, $p = 0.025$), indicating that hedonic consumption values strengthened the positive effect of professionalism on perceived value. Second, the interaction between authenticity and hedonic consumption values was negative and significant ($B = -0.107$, $p = 0.012$), indicating that hedonic consumption values weakened the positive effect of authenticity on perceived value. No significant interaction effect was found between practicality and hedonic consumption values.

To illustrate the two significant interaction effects, Figure 2 illustrates the two significant simple-slope patterns. For respondents with stronger hedonic consumption values, professionalism produced a steeper positive increase in perceived value. By contrast, the positive effect of authenticity on perceived value became weaker at higher levels of hedonic consumption values.

Table 11. Exploratory moderation results

Interaction Term	Dependent Variable	B	SE	Beta	p
Professionalism × Utilitarian Values	Perceived Value	-0.048	0.069	-0.054	0.487
Authenticity × Utilitarian Values	Perceived Value	0.107	0.058	0.118	0.069
Practicality × Utilitarian Values	Perceived Value	-0.085	0.055	-0.101	0.121
Professionalism × Hedonic Values	Perceived Value	0.126	0.056	0.115	0.025
Authenticity × Hedonic Values	Perceived Value	-0.107	0.042	-0.115	0.012
Practicality × Hedonic Values	Perceived Value	-0.039	0.048	-0.039	0.413
Professionalism × Utilitarian Values	Perceived Risk	-0.242	0.152	-0.195	0.112
Authenticity × Utilitarian Values	Perceived Risk	0.246	0.128	0.197	0.056
Practicality × Utilitarian Values	Perceived Risk	-0.180	0.121	-0.153	0.136
Professionalism × Hedonic Values	Perceived Risk	-0.134	0.123	-0.089	0.275
Authenticity × Hedonic Values	Perceived Risk	-0.113	0.093	-0.088	0.225
Practicality × Hedonic Values	Perceived Risk	0.198	0.105	0.142	0.059

Figure 2. Simple slope plots for the significant moderation effects of hedonic consumption values.



Taken together, the results indicate that the proposed framework was partially supported. The strongest and most stable findings concerned the direct effects of content characteristics and the mediating role of perceived value. By contrast, perceived risk was relatively weak in this context, and moderation effects emerged only in exploratory analysis involving hedonic consumption values.

4. Discussion

4.1 Beauty Recommendation Content Characteristics Matter Beyond Generic Video Attributes

The present study examined how three content characteristics of beauty recommendation videos - professionalism, authenticity, and practicality - influence college students' purchase decisions. Drawing on the stimulus-organism-response framework, the study compared two psychological pathways, perceived value and perceived risk, and further explored whether consumption values shaped selected effects. Overall, the findings suggest that beauty recommendation content affects college students mainly by enhancing perceived value rather than by reducing perceived risk. These findings provide a more

fine-grained explanation of short-form beauty video persuasion in social commerce.

A first important finding is that professionalism, authenticity, and practicality all positively influenced purchase decision. This result supports the view that short-form video persuasion is not driven only by platform exposure or general entertainment value. Instead, viewers actively evaluate the informational and experiential quality of the content before forming a decision tendency. This conclusion is broadly consistent with prior influencer marketing research showing that message value, credibility, and authenticity affect trust and purchase intention^[1]. It also aligns with short-form video studies that have emphasized usefulness, engagement, and creator-related perceptions as important predictors of consumer response^[4].

At the same time, the present study adds a more context-specific perspective. Much of the existing literature treats short-form video content through broad dimensions such as informativeness, ease of use, entertainment, or interaction^[7]. These dimensions are important, but they do not fully capture the decision logic of beauty recommendation content. Beauty consumption is strongly tied to personal suitability, usage experience, visual performance, and trial-based evaluation. In this context, professionalism, authenticity, and practicality are particularly meaningful because they address three key consumer concerns: whether the creator seems knowledgeable, whether the recommendation feels believable, and whether the content helps solve an actual purchase problem.

This refinement is theoretically useful because it suggests that short-form video content should not always be treated as a generic stimulus. In beauty social commerce, consumers may respond more strongly to domain-specific content cues than to broad platform-level attributes. The present findings therefore extend earlier studies by showing that beauty recommendation videos work not only as entertaining social media content, but also as decision-supporting information carriers.

4.2 Perceived Value Is the Core Psychological Mechanism

One of the most important findings of this study is that perceived value played a significant mediating role, whereas perceived risk did not. More specifically, professionalism, authenticity, and practicality all increased perceived value, and perceived value strongly predicted purchase decision. This indicates that beauty recommendation videos shape decision-making mainly by helping viewers feel that the recommended product is worthwhile, suitable, and potentially beneficial.

This result is consistent with previous research showing that trust and value-related evaluations are central mechanisms in influencer and short-form video marketing^[1]. Luo et al. found that content quality on short-video platforms promotes purchase intention through trust and perceived value^[4]. Likewise, prior beauty influencer studies have shown that viewers are more likely to respond positively when creators appear credible and when content supports trust-building^[3]. However, the present study extends those findings in an important way: rather than focusing primarily on trust, it highlights perceived value as the more direct and stable pathway linking beauty short-video content to purchase decision.

This value-centered mechanism is understandable in the beauty recommendation context. Beauty videos often show product texture, application effects, before-and-after comparisons, skin-fit suggestions, and price-related commentary. These cues make it easier for viewers to judge whether a product is worth trying or worth buying. As a result, the video does not merely create interest; it also supports value assessment. In other words, the content acts as a form of informal decision assistance, helping viewers estimate potential functional, emotional, and monetary returns.

Another notable point is that authenticity showed a particularly strong effect on perceived value. This suggests that in beauty recommendation settings, value is not built only through product claims or professional explanation. It is also built through a sense of realism. When the audience feels that the creator is honestly displaying product effects and limitations, the recommendation appears more diagnostic and therefore more valuable. This interpretation is broadly compatible with authenticity-based persuasion findings in influencer research^[2], but here the role of authenticity is linked specifically to value formation.

4.3 Why Perceived Risk Was Weak in This Context

In contrast to perceived value, perceived risk showed a much weaker role in the model. Only authenticity significantly reduced perceived risk, and perceived risk did not significantly predict purchase decision. As a result, none of the indirect paths through perceived risk were supported. This finding suggests that the psychological mechanism in this context is not primarily about risk reduction.

At a theoretical level, this result does not mean that perceived risk is irrelevant in consumer behavior. Rather, it suggests that in the specific setting of beauty recommendation videos for college students, risk concerns may be less central than value judgments. Several explanations are plausible.

First, many beauty products promoted in short-form videos are low- to mid-priced, highly replaceable, and easy to trial. In such cases, the perceived cost of making a wrong decision may be relatively limited. Consumers may therefore be more willing to experiment, especially when compared with high-priced durable goods or high-commitment service purchases. Under these conditions, the key question becomes not "What could go wrong?" but "Is this worth buying?" This helps explain why perceived value had a much stronger role than perceived risk.

Second, short-form beauty videos often reduce uncertainty by making product usage highly visible. Demonstrations, product comparisons, skin-type advice, and viewer comments may partially substitute for the lack of direct product trial. Although such cues may not eliminate all uncertainty, they can reduce ambiguity enough that risk ceases to be the dominant psychological barrier. This interpretation is consistent with the broader short-form video literature, which suggests that rich media content can strengthen user confidence and facilitate purchase-related judgments^{[4][6]}.

Third, the sample itself may matter. College students are intensive digital media users and are accustomed to processing large amounts of online recommendation content. They may rely on repeated exposure, peer comparison, and multiple creators' opinions to inform decisions. In that case, the influence of any single perceived risk cue may become diluted, while cumulative value impressions remain strong. This also suggests that consumer experience with digital content environments may weaken the explanatory power of traditional risk-based mechanisms in some social commerce settings.

From a contribution perspective, this finding is particularly useful because many studies tend to assume that trust enhancement and risk reduction are symmetric mechanisms. The present results suggest that these two mechanisms are not necessarily equal in explanatory power. In low-commitment beauty recommendation settings, value-building may be the more meaningful explanation. Future research should test whether the same pattern holds for high-priced, high-involvement, or safety-sensitive product categories.

4.4 Hedonic Consumption Values Play a Selective Rather Than Uniform Role

The moderation analysis produced a mixed but theoretically interesting pattern. The originally specified moderation assumptions were largely unsupported, especially for utilitarian consumption values and for perceived risk. However, additional interaction tests showed that hedonic consumption values strengthened the positive effect of professionalism on perceived value while weakening the effect of authenticity on perceived value.

This selective pattern suggests that hedonic orientation is not a general amplifier of all persuasive content. Instead, it changes how viewers interpret specific content cues. For respondents with stronger hedonic consumption values, professionalism may not be experienced only as technical expertise or factual guidance. In beauty consumption, professional content can also enhance the imagination of self-improvement, beauty upgrading, and aesthetic satisfaction. A creator who explains ingredients, techniques, or fit in a knowledgeable way may therefore make the product feel more aspirational and rewarding. Under such conditions, professionalism becomes not only informationally useful but also emotionally and symbolically valuable.

The negative interaction between hedonic consumption values and authenticity is more nuanced. One possible interpretation is that viewers with stronger hedonic orientations pay attention not only to realism but also to atmosphere, visual pleasure, and fantasy. In beauty short videos, highly authentic content may sometimes reduce the degree of aesthetic idealization that hedonic viewers enjoy. This does not mean that authenticity becomes unimportant. Rather, it suggests that for highly hedonic viewers, authenticity may not increase perceived value as strongly as it does for viewers who are less driven by pleasure and image-based enjoyment.

This result resonates partly with influencer marketing research showing that consumers respond differently to authenticity, relational closeness, and content style depending on their psychological orientation^{[2][7]}. Still, the present study adds a new nuance by showing that hedonic consumption values do not affect all beauty content dimensions in the same way. This has both theoretical and practical relevance because it suggests that consumer value orientation should be treated as a selective

boundary condition rather than a universal moderator.

4.5 Theoretical and Practical Implications

This study offers several theoretical implications. First, it extends the application of the stimulus-organism-response framework to a more specific short-form video social commerce context. Previous short-video studies have often examined general content features or platform experiences^{[4][5]}. By focusing on beauty recommendation videos, the present study shows that the SOR model remains useful when the stimulus is defined through more domain-specific content characteristics.

Second, the study refines the conceptualization of short-form video content. Instead of relying only on generic attributes such as entertainment or informativeness, it identifies professionalism, authenticity, and practicality as a meaningful content structure in beauty recommendation settings. This helps connect influencer marketing research with product-category-specific consumer logic.

Third, the study contributes to the literature on psychological mechanisms in digital persuasion by comparing perceived value and perceived risk within the same model. The results show that the two mechanisms do not operate with equal force. In this context, value-building is more central than risk-reduction. This distinction may help future studies better explain why some forms of social media persuasion succeed even when objective uncertainty is not fully removed.

Finally, the exploratory moderation results suggest that consumption values deserve more nuanced treatment in future research. Rather than assuming that utilitarian and hedonic values simply strengthen or weaken content effects in a uniform way, future studies should examine how different value orientations interact with specific content cues and product categories. The findings also provide several practical implications for brands, creators, and platforms. For beauty brands, the results suggest that content strategies should focus on increasing consumers' perceived value rather than only attempting to minimize perceived risk. In practical terms, this means emphasizing realistic product fit, usage guidance, and value-for-money cues. For creators, the results highlight the importance of balancing expertise and realism. Professional explanations and practical demonstrations are both effective, while authenticity remains a key driver of perceived value. For platforms, the results indicate that recommendation systems should not privilege only visual attractiveness or engagement metrics. Greater support for content that combines professionalism, authenticity, and practicality may improve not only user trust but also decision quality.

5. Conclusion

This study examined how beauty recommendation videos shape college students' purchase decisions through specific content characteristics. Drawing on the stimulus-organism-response framework, the study compared the mediating roles of perceived value and perceived risk and explored the boundary role of consumption values.

Three main conclusions can be drawn from the findings. First, professionalism, authenticity, and practicality all positively influence purchase decision, suggesting that college students do not respond to beauty recommendation videos only at a superficial or entertainment level. Instead, they evaluate whether the content is knowledgeable, believable, and useful for actual purchase decisions. Second, perceived value serves as the core psychological mechanism linking content characteristics to purchase decision, whereas perceived risk does not show a significant mediating role. This indicates that in the beauty recommendation context, value-building is more influential than risk-reduction. Third, hedonic consumption values do not uniformly strengthen all persuasive effects. Rather, they selectively shape how specific content cues are translated into value perception.

Overall, this study contributes to the literature in three ways. It refines the understanding of beauty-related short-form video content by identifying professionalism, authenticity, and practicality as distinct and meaningful dimensions. It extends the stimulus-organism-response framework to a beauty-focused short-video social commerce setting. It also shows that in a low-to medium-risk consumer context, perceived value may be more central than perceived risk in explaining digital persuasion outcomes.

Despite these contributions, the study has several limitations. The sample focused only on college students, which limits the generalizability of the findings to other consumer groups. Because the data were collected through a cross-sectional questionnaire, causal inferences should be made with caution. In addition, the study centered on beauty recommendation

videos in the Chinese context; future research may compare different platforms, product categories, or age groups. Further studies may also integrate additional psychological mechanisms, such as trust, parasocial interaction, or flow, to deepen the understanding of short-form video persuasion.

Funding

No

Conflict of Interests

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

Reference

- [1] Lou, C., & Yuan, S. (2019). Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. *Journal of Interactive Advertising*, 19(1), 58-73. <https://doi.org/10.1080/15252019.2018.1533501>
- [2] Liu, X., & Zheng, X. (2024). The persuasive power of social media influencers in brand credibility and purchase intention. *Humanities and Social Sciences Communications*, 11(15). <https://doi.org/10.1057/s41599-023-02512-1>
- [3] Garg, M., & Bakshi, A. (2024). Exploring the impact of beauty vloggers' credible attributes, parasocial interaction, and trust on consumer purchase intention in influencer marketing. *Humanities and Social Sciences Communications*, 11(235). <https://doi.org/10.1057/s41599-024-02760-9>
- [4] Luo, N., Pan, H., Mao, C., & Zhu, G. (2025). Influence of short video content on consumers' purchase intentions on social media platforms with trust as a mediator. *Scientific Reports*, 15, 16605. <https://doi.org/10.1038/s41598-025-94994-z>
- [5] Shen, X., & Wang, J. (2024). How short video marketing influences purchase intention in social commerce: The role of users' persona perception, shared values, and individual-level factors. *Humanities and Social Sciences Communications*, 11(290). <https://doi.org/10.1057/s41599-024-02808-w>
- [6] Zheng, C. (2023). Research on the flow experience and social influences of users of short online videos: A case study of Douyin. *Scientific Reports*, 13, 3312. <https://doi.org/10.1038/s41598-023-30525-y>
- [7] Sardar, S., & Vijay, T. S. (2025). Social media influencers and purchase intention: A review and research agenda. *International Journal of Consumer Studies*, 49(2), 233-246. <https://doi.org/10.1111/ijcs.70046>
- [8] Mehrabian, A. (2017). *Nonverbal communication*. Routledge. <https://doi.org/10.4324/9781315277349-5>
- [9] Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, 52(3), 2-22. <https://doi.org/10.1177/002224298805200302>
- [10] Sweeney, J. C., & Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing*, 77(2), 203-220. [https://doi.org/10.1016/S0022-4359\(01\)00041-0](https://doi.org/10.1016/S0022-4359(01)00041-0)
- [11] Stone, R. N., & Grønhaug, K. (1993). Perceived risk: Further considerations for the marketing discipline. *European Journal of Marketing*, 27(3), 39-50. <https://doi.org/10.1108/03090569310026637>
- [12] Babin, B. J., Darden, W. R., & Griffin, M. (1994). Work and/or fun: Measuring hedonic and utilitarian shopping value. *Journal of Consumer Research*, 20(4), 644-656. <https://doi.org/10.1086/209376>

Hybrid Course Design for Applied Management Learning in Higher Vocational Colleges: Digital Capability and Student Engagement as Pedagogical Pathways

Baoyin Liu, Rozaini Binti Rosli*

School of Business & Management, Lincoln University College, Petaling Jaya 47301, Selangor, Malaysia

**Corresponding author: Rozaini Binti Rosli, rozaini@lincoln.edu.my*

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: Introductory management courses in higher vocational colleges are expected to teach concepts, prepare students for workplace coordination, and cultivate practical judgement. Hybrid course design is often proposed as a remedy for the limits of lecture-centred delivery, yet its educational value depends less on the existence of an online platform than on how digital and face-to-face activities are sequenced. This article develops an integrative conceptual review of hybrid pedagogy in vocationally oriented management education. Drawing on literature on blended learning, constructive alignment, experiential learning, digital competence, and student engagement, it identifies two pedagogical pathways through which hybrid design may strengthen applied management learning: the development of digital capability and the formation of multidimensional engagement. Digital capability is treated as students' capacity to use digital tools for inquiry, collaboration, evidence handling, communication, and reflective task completion. Student engagement is understood as behavioral, cognitive, emotional, social, and agentic participation in learning. The article proposes a course-level framework in which well-designed online preparation, in-class problem work, collaborative case analysis, feedback, and assessment alignment combine to support applied learning outcomes. The review suggests that hybrid management courses are most effective when technology functions as a scaffold for practice rather than as a repository for content.

Keywords: Hybrid Course Design; Blended Learning; Management Education; Higher Vocational Colleges; Digital Capability; Student Engagement; Applied Learning Outcomes

Published: Apr 26, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1345>

1. Introduction

Introductory management courses occupy an awkward but important position in higher vocational colleges. They are usually placed early in a programme, so students encounter them before they have a mature vocabulary for organisations, markets, or managerial work. At the same time, these courses are expected to support employability: students must learn to read workplace situations, communicate with peers, explain decisions, and connect management concepts to concrete tasks. A course that only transmits definitions of planning, organising, leadership, and control may therefore satisfy the syllabus while leaving students unable to interpret a simple service failure, a delayed production schedule, or a team conflict. This gap between conceptual coverage and applied performance is especially visible in vocational settings, where the legitimacy of a management subject depends on whether students can see its usefulness in practice.

Hybrid or blended learning is often introduced as a response to this difficulty. In its broadest sense, blended learning combines online and face-to-face learning experiences in a planned educational arrangement^[1]. Earlier work has argued that such arrangements can be transformative when they are used to reconfigure teaching presence, learner responsibility, and the use of classroom time^[2]. More recent discussion cautions that the term has become too broad to be meaningful unless it is tied to specific design decisions^[3]. For introductory management education, the essential issue is not whether a course uses a platform, videos, quizzes, or discussion boards. The issue is whether these elements help students move from passive reception to disciplined management thinking.

A stronger rationale for hybrid management teaching emerges when the course is viewed through the logic of constructive alignment. Learning outcomes, activities, and assessment tasks should point toward the same forms of capability^[4]. In a vocational management course, such alignment means that students should not merely remember management theories; they should use concepts to diagnose situations, justify alternatives, and communicate feasible actions. Experiential learning theory offers a complementary perspective because it treats learning as a cycle of experience, reflection, conceptualisation, and testing^[5]. Hybrid teaching can support this cycle by distributing different moments across modes. Online spaces may support preparation, rehearsal, evidence search, and reflection, while face-to-face sessions may support debate, coaching, role play, and immediate clarification.

However, hybrid teaching may also reproduce the weaknesses of conventional teaching. A teacher may upload long slides, add a discussion forum that students rarely use, and then continue lecturing in class. In such cases, the online component merely increases workload without changing the quality of learning. The problem is partly technological, but it is also pedagogical. Teachers require the ability to connect content, teaching strategy, and technology in ways that serve disciplinary purposes, a concern that is consistent with the technological pedagogical content knowledge perspective^[6]. Management teachers need to know not only which platform functions are available, but also which digital routines help students interpret evidence, formulate arguments, and collaborate on managerial problems.

The course level is also a useful scale because it is where students experience pedagogy most directly. Policy documents may encourage digital transformation, and institutions may invest in platforms, but students encounter these decisions as weekly tasks, deadlines, discussions, and assessments. A framework that remains at the level of general strategy therefore gives teachers little guidance. By focusing on the learning sequence, the present review translates broad claims about hybrid learning into decisions that can be observed and revised within one semester. It also leaves room for local adaptation, which is essential in vocational programmes with different student profiles and industry links.

This article therefore asks a focused conceptual question: how can hybrid course design improve applied learning in introductory management courses in higher vocational colleges? It argues that two mechanisms are especially important. The first is digital capability, defined here as students' situated ability to use digital tools critically and purposefully for learning and work-like tasks. The second is student engagement, understood as a multidimensional pattern of participation rather than simple attendance or platform activity. The contribution of the article is a course-level framework that links hybrid design features to applied management outcomes through these two pathways. The article does not claim that hybrid delivery is automatically superior to face-to-face instruction. Its claim is more specific: hybrid pedagogy becomes educationally valuable when digital preparation, classroom interaction, feedback, and assessment are integrated around the practical uses of management knowledge.

2. Materials and Methods

2.1 Review Scope and Source Selection

The study adopts an integrative conceptual review design. This design is suitable where a field is shaped by several partially connected literatures rather than by a single stable empirical tradition. The review drew on five bodies of scholarship: blended and hybrid learning, management and business education, digital competence, student engagement, and vocationally oriented pedagogy. The aim was not to produce a statistically exhaustive mapping of every publication on blended learning. Instead, the review sought to identify concepts and design principles that can explain why hybrid formats may influence learning in introductory management courses.

Sources were selected through three criteria. First, priority was given to peer-reviewed articles, books, and reports that are widely used in higher education research. Second, selected sources had to offer concepts that could be translated into course design, such as learning alignment, teaching presence, engagement, feedback, or digital competence. Third, the source had to be relevant to practice-oriented learning rather than only to content delivery. For example, work on digital competence was included when it described the use of digital tools for information, communication, collaboration, or problem solving^{[8][9]}. Work on student engagement was included when it moved beyond the measurement of participation and addressed the quality of student involvement^{[10][11]}.

2.2 Synthesis Procedure

The synthesis followed a mechanism-based logic. The first step was to identify recurrent design features of effective hybrid teaching: concise preparatory input, low-stakes diagnostic tasks, face-to-face application, collaborative work, feedback loops, and reflective consolidation. The second step was to connect these features with learning processes in management education. Management courses require students to read ambiguous situations and make judgements under incomplete information; therefore, learning processes that promote inquiry, discussion, evidence handling, and reflection were treated as central. The third step was to interpret digital capability and engagement as mediating pathways between design and outcomes. This interpretation is consistent with the Community of Inquiry view that meaningful online and blended learning depends on the interaction of teaching, social, and cognitive presence^[7].

The analysis was deliberately bounded at the course level. Institutional strategy, learning management systems, and national policy matter, but the immediate concern of this article is how an instructor can design a basic management course that uses online and face-to-face modes coherently. The unit of analysis is therefore the recurring learning sequence: what students do before class, what they do in class, what feedback they receive, and how the work is assessed. This focus makes the framework useful for teachers who must operate within existing timetables, class sizes, and institutional platforms.

2.3 Analytical Categories

Three analytical categories guided the review. The first category was hybrid sequencing, which refers to the deliberate distribution of learning activities across modes. The second was digital capability, which refers to the academic and work-related use of digital tools rather than to general technology familiarity. The third was student engagement, which refers to behavioral, cognitive, emotional, social, and agentic participation. These categories were used to examine how a hybrid management course can support applied outcomes such as case diagnosis, evidence-based explanation, teamwork, concise communication, and reflective decision making.

3. Results

3.1 Hybrid Management Pedagogy as Structured Alternation

The first result of the review is that hybrid management pedagogy should be understood as structured alternation rather than as the coexistence of online and offline materials. In poorly designed hybrid courses, students experience the online part as separate homework and the classroom as a repetition of what they have already seen. In well-designed courses, each mode gives the other a reason to exist. Online preparation builds a minimum conceptual base; classroom activity tests and extends that base; post-class work consolidates the judgement formed through discussion. The course becomes a sequence of purposeful transitions.

A typical sequence in an introductory management topic can be described as follows. Before class, students watch a short explanation of a concept such as span of control, motivation, service quality, or conflict management. They then complete a brief diagnostic task, such as classifying a workplace incident or identifying two possible causes of a management problem. The teacher uses the responses to identify misconceptions. In class, students apply the concept to a richer case, negotiate interpretations in groups, and receive feedback on the quality of their reasoning. After class, they revise a short memo or submit a reflection that connects the case to a future workplace role. The sequence does not depend on expensive technology. It depends on the disciplined relation between preparation, application, and feedback.

This finding matters because introductory courses can easily become overloaded. Teachers often feel obliged to cover many theories, while students struggle to distinguish core ideas from peripheral details. Hybrid design can either worsen or reduce

this problem. It worsens the problem when every lecture is supplemented by additional videos, readings, and quizzes without a clear purpose. It reduces the problem when online materials are concise and when class time is reserved for interpretive work that cannot be completed by solitary reading. Active learning research suggests that student learning improves when learners are required to process, discuss, and apply ideas rather than only listen^{[14][15]}. Hybrid design can create the conditions for such activity, but only when the online part prepares students for participation rather than replacing it.

Table 1. Structured hybrid sequence for an introductory management topic

Phase	Typical activity	Pedagogical function	Applied management value
Pre-class input	Short video, annotated slide deck, or guided reading on a core concept.	Creates a shared conceptual starting point.	Students enter class with the vocabulary needed for diagnosis.
Diagnostic task	Brief quiz, case classification, or one-paragraph response.	Reveals misconceptions before face-to-face teaching.	Teacher feedback can target errors in reasoning rather than repeat content.
Face-to-face application	Group case analysis, role play, problem-solving discussion, or decision memo workshop.	Turns abstract concepts into situated judgement.	Students practise communication, trade-off analysis, and collaborative reasoning.
Feedback loop	Peer review, teacher questioning, model answer comparison, or rubric-based comments.	Makes quality criteria visible.	Students learn what counts as a convincing management explanation.
Post-class consolidation	Revised memo, reflective note, or workplace transfer task.	Stabilises learning after discussion.	Students connect classroom cases to future work situations.

3.2 Digital Capability as an Enabling Pathway

The second result is that digital capability should be treated as an enabling pathway rather than as a background characteristic. Many students in higher vocational colleges use mobile devices frequently, but frequent use does not guarantee academic or professional digital competence. Students may be comfortable with messaging and social media while still struggling to search for credible information, organise evidence, collaborate in shared documents, or present a brief management recommendation. Digital competence frameworks emphasise information literacy, communication, content creation, safety, and problem solving^{[8][9]}. In management education, these domains need to be connected to workplace-like tasks.

A hybrid management course can cultivate digital capability when digital tools are used to support the practices of management work. Students can search for evidence on customer complaints, compare organisational charts, analyse simple survey responses, or prepare a short briefing for a simulated supervisor. They can also use collaborative documents to divide roles, track contributions, and produce group outputs. These activities are different from asking students to submit answers through a platform. They require students to use technology to inquire, coordinate, and communicate. In this sense, digital capability is not a separate graduate attribute placed beside management knowledge; it is a medium through which management knowledge is enacted.

The capability pathway also has an equity dimension. Hybrid learning may disadvantage students when it assumes that all learners have stable access, private study spaces, or the same level of digital confidence. The design response should not be to avoid digital tasks, because workplace environments increasingly expect digital coordination. Rather, teachers should scaffold the tasks. For instance, a first-year course can begin with a guided search template, a shared-document protocol, and explicit instructions on how to cite an online source. Later tasks can gradually reduce the scaffolding and require students to make more independent digital choices. In this way, the hybrid course develops capability over time instead of treating it as a precondition for success.

Digital capability is also closely related to credibility in management reasoning. A student who can locate basic evidence, distinguish an assertion from a supported claim, and present data clearly is better positioned to make a persuasive management argument. Introductory management courses often deal with familiar topics such as motivation, leadership,

teamwork, and service quality. Because these topics feel familiar, students may rely on personal opinion. Digital tasks can counter this tendency when they require students to support interpretations with evidence, compare sources, or visualise a simple pattern. The educational value lies not in the tool itself but in the disciplined use of the tool to improve judgement.

Table 2. Digital capability dimensions in vocational management learning

Dimension	Course task example	Management learning value
Information and evidence literacy	Find and compare evidence related to a service failure, employee complaint, or customer trend.	Improves evidence-based diagnosis and reduces opinion-only reasoning.
Digital communication	Prepare a concise briefing slide, recorded explanation, or written memo for a simulated manager.	Builds clarity, audience awareness, and professional expression.
Collaboration and coordination	Use a shared document to allocate roles, track progress, and assemble a group case response.	Connects teamwork theory with the experience of coordinating actual work.
Data handling	Summarise a small data set from a class survey or simple business scenario.	Supports practical interpretation of operational or customer information.
Reflective digital practice	Revise a decision memo after peer and teacher feedback.	Encourages students to treat feedback as part of professional improvement.

3.3 Student Engagement as the Proximal Pathway

The third result is that student engagement is the proximal pathway through which hybrid design influences learning. Engagement cannot be reduced to attendance, log-in frequency, or the completion of quizzes. Research commonly distinguishes behavioral, cognitive, and emotional engagement^[10], while later work emphasises the broader interaction between students and their educational environment^[11]. In hybrid management courses, social and agentic dimensions are also important. Students engage socially when they negotiate meaning with peers, and they engage agentially when they ask questions, propose examples, or influence the direction of a discussion.

Behavioral engagement appears when students prepare, attend, submit tasks, and contribute to group work. Cognitive engagement appears when students connect theories to cases, compare alternatives, and monitor the quality of their own explanations. Emotional engagement appears when students experience the course as relevant rather than as a list of abstract terms. Social engagement appears through collaboration, peer questioning, and shared responsibility for a group output. Agentic engagement appears when students bring workplace observations, part-time job experiences, or local business examples into the learning process. Hybrid design can support each of these forms, but only if the activities require visible participation.

Several design practices are particularly useful. Short pre-class tasks create a reason to prepare. In-class case work creates a reason to speak and listen. Shared digital documents create a record of contribution. Feedback creates a reason to revise. Assessment rubrics create a reason to care about the quality of explanation rather than only the correctness of an answer. These practices align with research on online engagement strategies, which emphasises instructor presence, structured interaction, and meaningful learning tasks^[12]. They also reflect findings from educational technology studies showing that engagement improves when technology is embedded in purposeful pedagogy rather than used as an isolated novelty^[13].

In vocational management education, engagement is inseparable from relevance. Students may not immediately recognise why a classical theory of motivation or organisational structure matters. Engagement improves when teachers frame such theories through realistic decisions: how to assign responsibilities in a small team, how to respond to a customer complaint, how to communicate a change in a work schedule, or how to reduce conflict between employees. Hybrid design helps because students can encounter the basic vocabulary before class and then spend face-to-face time on judgement. The course becomes less about remembering theory names and more about using theory as a tool for interpreting work.

3.4 Integrated Pathway Model and Applied Learning Outcomes

The review leads to an integrated pathway model. The first layer of the model is course design: online preparation, diagnostic tasks, in-class application, collaborative production, feedback, and assessment alignment. The second layer is the mediating process: digital capability and student engagement. The third layer is applied learning outcomes: case diagnosis, evidence-based explanation, teamwork, managerial communication, and reflective transfer. The central argument is that hybrid design affects outcomes indirectly. A course does not improve simply because content is distributed across digital and physical spaces. It improves when that distribution develops capabilities and increases the quality of students' participation.

The model also clarifies the role of feedback. Feedback is often treated as a comment given after a task, but in hybrid design it should be built into the learning sequence. A diagnostic quiz before class is a form of feedback to the teacher because it reveals what students have misunderstood. Peer discussion is feedback to students because it exposes them to alternative interpretations. A rubric is feedback before submission because it tells students what features of reasoning will be valued. Effective feedback helps learners answer three questions: what is the goal, where am I now, and what should I do next^[17]. In management courses, these questions translate into whether students can identify the problem, justify their diagnosis, and improve the practicality of their recommendation.

Assessment alignment is the final element of the model. If final assessment rewards memorisation alone, students will rationally treat online tasks and case discussions as peripheral. If assessment includes a case memo, an oral briefing, a group analysis, or a reflective application task, students have a clearer reason to engage with the hybrid sequence. This does not mean abandoning conceptual knowledge. It means assessing concepts through their use. A student can still be required to explain leadership, motivation, or control, but the explanation should be connected to a situation that demands interpretation. In this way, assessment confirms that hybrid learning is not an extra layer added to the course; it is the organising logic of the course.

Table 3. Design principles for hybrid management courses

Principle	Implementation	Risk if ignored
Integrate modes	Every online task should be used in class, feedback, or assessment.	Students treat online work as optional busywork.
Limit content load	Use concise preparatory materials and reserve class time for application.	Hybrid delivery becomes a heavier version of lecturing.
Make participation visible	Use diagnostic tasks, shared documents, and brief outputs.	Teachers cannot distinguish preparation from passive attendance.
Teach digital routines	Model searching, source evaluation, collaboration protocols, and memo revision.	Digital tasks reproduce inequality in confidence and access.
Assess applied reasoning	Use cases, briefings, memos, and reflective transfer tasks.	Students memorise terms without developing practical judgement.

4. Discussion

4.1 Course Design Implications

The framework has practical implications for teachers of basic management courses. First, teachers should design backward from applied outcomes. If the expected outcome is that students can analyse a workplace coordination problem, the course should not begin with a long lecture on every theory of organisation. It should begin with a manageable concept, a short case, and a task that requires students to use the concept. The online component should prepare students for this task, not duplicate the textbook. The face-to-face component should then require interpretation, argument, and feedback. Such design is consistent with constructive alignment because students practise the form of thinking that they will later be assessed on^[4].

Second, teachers should protect face-to-face time. The classroom is valuable because it allows students to hear disagreement, defend interpretations, observe modelling by the teacher, and receive immediate clarification. If the teacher uses class time mainly to repeat online content, the hybrid model loses its advantage. A more productive class session may begin with a

five-minute review of common errors from the pre-class task, followed by group analysis of a case, teacher questioning, and a brief written conclusion. This arrangement is not technologically complex, but it changes the centre of the lesson from exposition to guided practice.

Third, teachers should treat digital capability as part of the curriculum. This requires explicit instruction. Students can be shown how to search for a credible source, how to summarise evidence without copying, how to organise a shared document, and how to revise a memo after feedback. These routines may appear minor, but they are central to employability. In many workplaces, management communication is mediated by documents, platforms, dashboards, and asynchronous messages. A vocational management course that ignores these practices risks presenting management as a set of textbook definitions rather than as a communicative and evidence-based activity.

Fourth, teachers should be careful with workload. Hybrid learning can fail when students perceive it as double teaching: online lectures plus full classroom lectures plus extra assignments. The design principle should be substitution and integration rather than accumulation. A pre-class video can replace part of the lecture; a diagnostic task can replace a routine homework exercise; a revised memo can replace a conventional short-answer quiz. When students see that each task is used in class or assessment, they are more likely to interpret the hybrid structure as purposeful.

4.2 Assessment and Institutional Implications

Assessment is the point at which the credibility of hybrid teaching is tested. A hybrid course that claims to develop applied management learning should include assessment that captures applied reasoning. Possible formats include short case analyses, group decision briefings, reflective learning logs, oral explanations, or portfolio tasks. Formative assessment is especially useful because first-year vocational students may need repeated opportunities to see the difference between description and analysis. Research on formative assessment and feedback indicates that learners benefit when feedback supports self-regulation and improvement rather than only judgement after the fact^{[16][18]}.

Institutions also have responsibilities. Course-level design depends on platform reliability, teacher workload, classroom arrangements, and professional development. Teachers need time to redesign materials, not simply instructions to upload resources. They also need support in making digital tasks accessible to students with uneven devices or connectivity. Professional development should therefore move beyond technical training. It should help teachers convert a topic into a hybrid sequence, design a diagnostic task, read student responses, and facilitate a case discussion. This kind of support recognises that hybrid teaching is a pedagogical practice, not a software function.

The framework may also guide programme-level coordination. If every course uses a different platform routine, students spend unnecessary energy learning formats rather than learning management. Programmes can agree on a small number of common routines: short preparatory input, pre-class check, group case document, feedback rubric, and revised output. Such routines reduce confusion while still allowing each teacher to adapt content. They also make digital capability cumulative across courses. Students who learn to prepare a management memo in one course can later use the same routine for marketing, human resource management, entrepreneurship, or operations management.

4.3 Risks, Boundary Conditions, and Equity

The proposed framework should not be interpreted as a universal guarantee. Several boundary conditions are important. Large class size may limit the amount of individual feedback a teacher can provide. Weak platform access may reduce the reliability of pre-class tasks. Students with low academic confidence may avoid online discussion if the environment feels punitive. Teachers with limited design experience may find it easier to upload content than to build learning sequences. These conditions do not invalidate hybrid pedagogy, but they require design choices that are modest and sustainable. A course can begin with one weekly diagnostic task and one case discussion rather than attempting a complete redesign at once.

Equity requires particular attention. Hybrid teaching can support inclusion when online materials allow students to review content at their own pace, prepare before speaking, and receive feedback in multiple forms. It can also create exclusion when access problems, unclear instructions, or hidden digital expectations are ignored. Teachers should therefore provide alternative access routes where possible, use mobile-friendly materials, keep videos short, and explain how online work connects to class activity. The goal is not to lower expectations, but to make expectations transparent. Students should be

challenged to use digital tools professionally, while also receiving the scaffolding needed to reach that standard.

Another risk is superficial engagement. Students may complete quizzes mechanically, copy from online sources, or divide group work without genuine discussion. The response is to design tasks that require interpretation and traceable contribution. For instance, a group case document can include a section for evidence, a section for alternative explanations, and a section for the final recommendation. Peer review can ask students to comment on the quality of reasoning rather than simply agree or disagree. In this way, the digital record becomes a support for thinking and accountability, not merely a submission channel.

4.4 Limitations and Future Research

This article has limitations. It is an integrative conceptual review, not a systematic review or an empirical test. The framework therefore requires validation through classroom studies. Future research could examine whether hybrid sequences improve students' case analysis, communication quality, or reflective transfer compared with lecture-centred formats. Mixed-methods designs would be useful because quantitative measures of achievement can be combined with qualitative analysis of student memos, discussion patterns, and teacher feedback. Learning analytics may also help identify whether online preparation predicts the quality of in-class participation, but such data should be interpreted carefully and ethically.

Further research should also examine moderators. Prior digital capability, language proficiency, teacher design competence, class size, platform usability, and assessment culture may all affect the success of hybrid management courses. Studies in higher vocational colleges are particularly needed because much of the blended learning literature has been developed in universities or fully online environments. Vocational contexts have distinctive features: students may be more practically oriented, curricula may emphasise employability, and teachers may be expected to connect theory with industry scenarios. These features make the context valuable for developing a more situated account of hybrid pedagogy.

Finally, future research can refine the applied outcomes proposed in this article. General achievement scores may not capture the learning that matters in management education. More precise indicators could include the accuracy of problem diagnosis, the relevance of evidence, the clarity of a recommendation, the quality of teamwork, and the ability to revise work after feedback. These indicators would allow researchers to test whether digital capability and engagement truly mediate the relationship between hybrid design and learning outcomes. They would also give teachers more actionable evidence for redesigning their courses.

5. Conclusions

Hybrid pedagogy offers a promising but easily misunderstood route for improving introductory management courses in higher vocational colleges. Its value does not lie in transferring lectures to a platform or adding online tasks to an already crowded syllabus. Its value lies in creating a coherent learning sequence in which students prepare, apply, discuss, receive feedback, and revise. This article has argued that two pathways explain why such design can matter: the development of digital capability and the strengthening of multidimensional engagement. Digital capability enables students to use tools for evidence, communication, collaboration, and reflective improvement. Engagement gives those practices educational force by drawing students into behavioral, cognitive, emotional, social, and agentic participation.

The proposed framework places applied management learning at the centre of hybrid design. Students should leave an introductory management course not only with terms and theories, but also with the ability to use those theories when they interpret workplace problems. For teachers, the main task is to make the relation between online preparation, classroom activity, feedback, and assessment visible. For institutions, the main task is to provide the conditions under which such design is sustainable. When these conditions are met, hybrid management courses can help vocational students move from knowing about management to practising the early forms of managerial judgement that their future work will require.

Funding

No

Conflict of Interests

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

Reference

- [1] Graham, C. R. (2006). Blended learning systems: Definition, current trends, and future directions. In C. J. Bonk & C. R. Graham (Eds.), *The handbook of blended learning: Global perspectives, local designs* (pp. 3-21). Pfeiffer.
- [2] Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, 7(2), 95-105. <https://doi.org/10.1016/j.iheduc.2004.02.001>
- [3] Hrastinski, S. (2019). What do we mean by blended learning? *TechTrends*, 63, 564-569. <https://doi.org/10.1007/s11528-019-00375-5>
- [4] Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Open University Press.
- [5] Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.
- [6] Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054.
- [7] Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2-3), 87-105.
- [8] Redecker, C. (2017). European framework for the digital competence of educators: DigCompEdu. Publications Office of the European Union. <https://doi.org/10.2760/159770>
- [9] Vuorikari, R., Kluzer, S., & Punie, Y. (2022). DigComp 2.2: The digital competence framework for citizens. Publications Office of the European Union. <https://doi.org/10.2760/115376>
- [10] Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- [11] Kahu, E. R. (2013). Framing student engagement in higher education. *Studies in Higher Education*, 38(5), 758-773. <https://doi.org/10.1080/03075079.2011.598505>
- [12] Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, 22(1), 205-222. <https://doi.org/10.24059/olj.v22i1.1092>
- [13] Bond, M., Buntins, K., Bedenlier, S., et al. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education*, 17, 2. <https://doi.org/10.1186/s41239-019-0176-8>
- [14] Freeman, S., Eddy, S. L., McDonough, M., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415. <https://doi.org/10.1073/pnas.1319030111>
- [15] Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223-231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809.x>
- [16] Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. <https://doi.org/10.1080/03075070600572090>
- [17] Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- [18] Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712. <https://doi.org/10.1080/02602938.2012.691462>

Reconstructing Vocational Teacher Capacity in AI-Enabled Industry-Education Integration: A Policy and Institutional Analysis of Liaoning Province

Chi Tang, Rozaini Binti Rosli*

School of Business & Management, Lincoln University College, Petaling Jaya 47301, Selangor, Malaysia

*Corresponding author: Rozaini Binti Rosli, rozaini@lincoln.edu.my

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 digital transformation of vocational education has moved beyond the question of whether teachers should use digital tools. In regions where industrial renewal and vocational reform overlap, the more difficult question is how teachers' professional capacity should be rebuilt when artificial intelligence, data-based training resources and school-enterprise co-development become part of everyday institutional work. Taking Liaoning Province as the analytical setting, this paper examines how vocational teachers' roles are being reorganised under three connected pressures: the spread of AI-enabled pedagogy, the demand for deeper industry-education integration and the regional need to support industrial upgrading. The study adopts qualitative document analysis of international frameworks, Chinese national policy documents, Liaoning provincial education documents and recent empirical studies on TVET teacher digital competence. The findings suggest that teacher role change is better understood as capacity reconstruction than as simple role replacement. Four dimensions are identified: curriculum translation between occupational tasks and learning outcomes, AI-assisted instructional design and verification, boundary work in school-enterprise cooperation, and ethical stewardship of digital assessment and learner data. The paper argues that AI application strengthens vocational education only when it is embedded in industry-linked curriculum renewal and supported by organisational arrangements for teacher development. For vocational colleges in Liaoning, the practical priority is not to train teachers to operate isolated tools, but to build collaborative mechanisms through which teachers, enterprises and digital platforms jointly update curriculum, assessment and workplace learning. The paper offers a regional policy-informed framework for managing vocational teacher development in AI-enabled vocational education.

Keywords: Vocational Education; Artificial Intelligence; Teacher Capacity; Industry-Education Integration; School-Enterprise Co-Development; Digital Pedagogy; Liaoning Province

Published: Apr 26, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1346>

1. Introduction

Technical and vocational education and training has again become a strategic site for economic and social policy. It is expected to prepare learners for employment, support lifelong learning, absorb the pressure of industrial restructuring and, increasingly, help societies move towards digital and green transitions^[1]. In China, this expectation has been reinforced through the National Vocational Education Reform Implementation Plan and the revised Vocational Education Law, both of which reposition vocational education as a type of education with its own value rather than a secondary route below general

education^[2,3]. These policy shifts matter for teachers because the quality of vocational education is not determined only by equipment, platforms or enrolment scale. It depends on whether teachers can connect changing occupational standards with learnable curriculum, practical training and credible assessment.

Artificial intelligence has sharpened this issue. The rapid diffusion of generative AI, learning analytics, intelligent tutoring systems and simulation-based training has made it easier to produce lesson materials, demonstrations, feedback and practice tasks. Yet the same technologies also create new risks: plausible but inaccurate technical explanations, over-automated feedback, weak data protection, algorithmic bias and a widening gap between platform functions and actual workplace requirements. UNESCO's guidance on generative AI therefore stresses human-centred governance, while its AI competency framework for teachers treats AI literacy as a matter of professional judgement, ethics and pedagogy, not merely tool operation^[4,5]. For vocational education, these concerns are especially acute because teaching errors may affect not only examination performance but also occupational safety, production standards and learners' capacity to transfer knowledge into real work.

The teacher's role in vocational education has never been limited to classroom instruction. The OECD's review of teachers and leaders in vocational education describes VET teachers as a workforce located between pedagogical institutions and the labour market, requiring access to up-to-date occupational knowledge, practical experience and professional development^[6]. Empirical studies have reached a similar conclusion from the digital competence side. Research on Chinese TVET teachers has found that digital competence is influenced by both personal and institutional factors, indicating that teacher development cannot be reduced to individual willingness or age-related familiarity with technology^[7]. Studies in European contexts further show that vocational teachers' technology use depends on perceived usefulness, confidence, workload and the teaching profile in which the teacher works^[8,9]. Pre-service vocational teacher research also indicates that digital competence is multidisciplinary, combining technical knowledge, motivational factors, cognitive strategies and pedagogical application^[10].

This article takes Liaoning Province as a regional case because it sits at the intersection of vocational education reform and industrial renewal. Liaoning is a traditional industrial base with policy attention to equipment manufacturing, petrochemicals, shipbuilding, robotics, aviation, digital manufacturing and new service industries. Its education plan calls for the construction of digital vocational education, digital teaching resources, online open courses, vocational education data platforms and learning platforms, while also supporting modern apprenticeship, school-enterprise dual education and the incorporation of new technologies, processes and standards into teaching content^[11]. More recent provincial education materials report the establishment of industry-education communities in fields such as petrochemicals, robotics, shipbuilding and aviation, together with municipal industry-education alliances, school-enterprise training bases and projects linked to regional industrial needs^[12]. These developments make Liaoning a useful setting for examining how teacher roles are being reshaped by the combined force of AI application and school-enterprise co-development.

The paper deliberately avoids treating AI as an autonomous force that simply changes teachers from outside. Such a view overstates the technology and understates the institutional work required to make digital transformation meaningful. In vocational colleges, AI can draft, recommend, classify, simulate and assess, but it cannot by itself decide which occupational tasks are worth teaching, which standards should be translated into learning outcomes, or how students' weak performance in a simulated task should be interpreted. These tasks still require teachers who understand industry practice, learners' development and institutional constraints. At the same time, the traditional idea of the vocational teacher as a stable transmitter of occupational knowledge is also insufficient. When technologies and production processes change quickly, teachers must become interpreters, designers, validators and coordinators.

From an economic and management perspective, this issue should be understood as a problem of regional human-capital governance. Liaoning's vocational colleges are not only education providers; they are organisations through which industrial policy, enterprise demand and labour-market transition are translated into trainable capabilities. Teacher capacity is thus a managerial asset. When it is weak, digital investment may remain at the level of platform purchase. When it is strong, the same investment can support curriculum updating, learner employability and enterprise-facing innovation. The teacher becomes a site where regional development strategy is operationalised.

The guiding research question is therefore: how should vocational teachers' professional capacity be reconstructed when AI application and industry collaboration are treated as mutually connected conditions of digital transformation? The article answers this question through a qualitative analysis of policy documents and research literature. It contributes a framework that links teacher capacity reconstruction to four forms of work: translating occupational change into curriculum, designing and checking AI-supported learning, brokering school-enterprise knowledge flows, and safeguarding ethics and learner development. Rather than asking whether teachers will be replaced by AI, the article asks what kinds of judgement and organisational support are needed for teachers to use AI responsibly within vocational education.

2. Materials and Methods

2.1 Research Design

The study uses qualitative document analysis. This method is appropriate because the transformation of vocational education is being advanced through policies, standards, institutional plans and professional development schemes before it becomes fully visible in routine classroom practice. Document analysis allows the researcher to examine how problems are defined, how responsibilities are allocated and how specific forms of professional capacity are made thinkable within policy discourse^[13]. To strengthen interpretive discipline, the analysis was guided by content-analytic principles: documents were read for recurring categories, explicit policy requirements, implied role expectations and tensions between different levels of governance^[14]. The purpose was not to measure teachers' actual behaviour in a statistical sense, but to reconstruct the professional logic that current policy and research attach to vocational teachers under AI-enabled digital transformation.

2.2 Document Corpus

The document corpus included four types of materials. The first group consisted of international frameworks on TVET, AI in education and vocational teacher development, including UNESCO and OECD publications. These texts helped establish the broader conceptual vocabulary concerning human-centred AI, teacher agency and work-based relevance. The second group consisted of Chinese national policy documents on vocational education, teacher digital literacy and AI-related educational transformation. The third group consisted of Liaoning provincial education documents, including the provincial education development plan and public replies or reports concerning vocational education, teacher development, school-enterprise cooperation, professional capability mapping and digital-intelligence literacy. The fourth group consisted of peer-reviewed empirical studies on vocational teachers' digital competence, technology use and boundary-crossing identity. This mixed corpus was selected because teacher role reconstruction is simultaneously a policy, institutional and professional phenomenon.

The inclusion criteria were relevance to vocational education, teacher work, digital transformation, AI application, school-enterprise cooperation or Liaoning's vocational education reform. Documents were excluded when they discussed general digital economy policy without an education component, when they dealt only with basic education, or when they made claims about AI without connecting them to teacher work or institutional governance. Although the corpus did not seek exhaustive coverage of all documents published in China, it was designed to capture the main policy and research signals shaping vocational teacher capacity in the Liaoning context.

2.3 Analytical Procedure

The analysis proceeded in three rounds. In the first round, all documents were coded for explicit references to teachers, AI, digital resources, curriculum reform, workplace training, school-enterprise cooperation and assessment. In the second round, these codes were grouped into broader categories of teacher work: curriculum design, technological application, industry translation, learner support, assessment and ethical governance. In the third round, the categories were compared across international, national and Liaoning-level documents. This comparison was important because international frameworks tend to emphasise human agency and ethics, national documents tend to emphasise system-building and standards, and provincial documents tend to emphasise implementation, industrial fit and training mechanisms. The four findings reported below were derived from convergence and tension across these layers.

Two precautions were taken to avoid over-interpretation. First, the study distinguishes between policy intention and classroom reality. A policy document may require school-enterprise cooperation or AI literacy training, but such requirements do not

automatically prove that teachers already possess the necessary competence. Second, the analysis treats Liaoning as a policy-informed regional case, not as a statistically representative sample of all Chinese provinces. The argument is strongest where it identifies institutional pathways and professional demands; it does not claim to measure the prevalence of those pathways among all vocational teachers.

3. Results

3.1 From Technical Updating to Capacity Reconstruction

The first finding is that vocational teacher development is moving from technical updating towards broader capacity reconstruction. In older forms of in-service training, digital transformation could be understood as adding a new platform or teaching software to an otherwise stable curriculum. Current documents imply a different situation. Liaoning's education plan connects digital vocational education with professional digital upgrading, online open courses, data platforms, student-centred classroom reform and the inclusion of new technologies, processes and standards in teaching content^[11]. This combination is significant. It suggests that teachers are not merely expected to digitise existing lessons; they are expected to rethink what should be taught and how occupational knowledge should be organised when industries themselves are changing.

Capacity reconstruction therefore has at least three layers. The first is operational: teachers must know how to use digital platforms, simulation environments, AI tools and online resources. The second is pedagogical: teachers must convert these tools into learning designs that support practice, reflection and transfer. The third is occupational: teachers must maintain enough contact with industry to understand how digital systems are actually used in workplaces. A teacher who can operate a platform but cannot judge whether a generated case reflects current production practice has only partial digital competence. Conversely, a teacher with rich occupational experience but little capacity to design hybrid or AI-supported learning will find it difficult to help students practise in new digital environments.

The revised understanding of capacity also changes the meaning of the double-qualified teacher. In many vocational systems, the double-qualified teacher has been described as someone with both teaching competence and occupational competence. Under AI-enabled digital transformation, this duality becomes more complex. Teachers are required to connect pedagogical knowledge, occupational knowledge and digital judgement. Liaoning's public response on improving professional-course teachers' skills highlights the role of enterprises in teacher development, naming industry organisations and firms that participate in training design, course teaching and practical operation so that production and technical resources can be opened to vocational education^[15]. This is not only a mechanism for updating skills. It is a way of making teachers' occupational knowledge current enough to support digital curriculum renewal.

The policy implication is that the target of teacher development should not be a generic digitally competent teacher, but a teacher able to handle the relationship between digital tools and occupational change. This distinction matters for college management. If professional development is defined narrowly as software training, teachers may learn to produce more digital materials without improving the relevance of student learning. If it is defined as capacity reconstruction, then the training agenda must include work-task analysis, curriculum redesign, AI-assisted resource evaluation, digital assessment, collaborative lesson study and periodic industry immersion. This broader agenda is more demanding, but it better matches the kind of teacher work now required in vocational education.

3.2 AI Application and the Teacher as Designer, Validator and Steward

The second finding concerns AI application. The documents and literature do not support a simple replacement narrative. AI can assist with lesson preparation, resource generation, student feedback, formative assessment, simulation and administrative work, but these functions increase rather than remove the need for teacher judgement. UNESCO's framework identifies human-centred mindset, ethics, AI foundations, AI pedagogy and professional learning as dimensions of teacher AI competence^[5]. In vocational education, each dimension has a practical form. Teachers must know when AI-generated explanations are technically reliable, when a simulation oversimplifies workplace conditions, when automated feedback misses a safety-critical error, and when student data should not be used for certain forms of prediction.

The teacher's design role becomes more visible when AI is used in practical training. A vocational teacher may ask an AI system to generate a troubleshooting scenario for a manufacturing line, a customer service dialogue, a logistics routing

problem or a nursing communication exercise. The output may be useful, but it still requires professional checking. The teacher must judge whether the task reflects the right level of difficulty, whether the terminology is aligned with current industry usage, whether the case encourages procedural compliance or deeper problem-solving, and whether assessment criteria are transparent. In this sense, AI shifts part of the teacher's work from producing every resource manually to curating, adapting and validating resources. The labour is different, not absent.

The validation role is particularly important in vocational education because many learning tasks carry embedded safety and quality standards. A language error in a general essay prompt may be inconvenient; a misleading maintenance procedure, chemical handling step or machine operation instruction can be dangerous. AI systems are also prone to presenting answers with a degree of fluency that can hide uncertainty. Teachers must therefore teach students not only how to obtain AI-generated information, but how to interrogate it. This is a curricular issue. Students in vocational programmes need habits of verification: checking against technical manuals, comparing with enterprise standards, consulting experienced workers, and explaining why a proposed solution is acceptable in a given workplace context.

The stewardship role is equally important. AI-supported teaching can generate data on student progress, mistakes, response time and behavioural patterns. Used carefully, such data can help teachers identify weak skills and personalise practice. Used carelessly, it can narrow learning to what is easily measurable, stigmatise learners or create opaque forms of surveillance. The Ministry of Education's Teacher Digital Literacy standard emphasises teachers' awareness, ability and responsibility in using digital technology to optimise and transform teaching activities^[16]. The 2026 AI + Education Action Plan goes further by calling for teacher AI literacy standards, classified training and assessment systems, as well as AI support for lesson preparation, learning analysis and digital resource generation^[17]. These policy signals imply that AI-related teacher competence must include ethical restraint and data governance.

A practical consequence is that vocational colleges should not evaluate AI use by counting how many teachers have used a tool. A more meaningful evaluation would ask whether AI use improves occupationally authentic learning, whether generated materials are checked by teachers and enterprises, whether students learn verification practices, and whether data use is transparent. The teacher remains the professional agent who decides what counts as good evidence of learning. AI can expand the evidence base, but it cannot replace the normative judgement embedded in vocational education.

3.3 School-Enterprise Co-Development and Teacher Boundary Work

The third finding is that industry collaboration mediates AI-enabled teacher role reconstruction. Digital transformation in vocational education becomes superficial if it is confined to campus platforms. Occupations are transformed in enterprises, workshops, service systems and production networks. Teachers therefore need channels through which workplace changes enter curriculum. Recent vocational teacher literature describes such work as boundary crossing: teachers move between educational and occupational communities, and their professional identity is formed through negotiating these different expectations^[18]. Under AI conditions, the boundary becomes more complicated because digital tools can simulate work without guaranteeing that the simulation reflects actual workplace practice.

Liaoning's policy materials show that school-enterprise cooperation is no longer framed only as internship placement. The province has reported industry-education communities in petrochemicals, robotics, shipbuilding and aviation, municipal industry-education alliances, 555 school-enterprise practice and training bases, and project connections among 617 member units^[12]. It has also reported efforts to match enterprise post requirements with vocational education resources through modern apprenticeship and on-site engineer training, along with the construction of professional capability maps based on regional industrial needs^[19]. These initiatives indicate that the relevant unit of collaboration is not a single enterprise visit but a network of curriculum, training base, capability model and employment demand.

For teachers, this creates a boundary-broker role. They must translate enterprise needs into educational tasks, and translate educational limits back to enterprises. Enterprises may want graduates who can immediately operate a system; colleges must also build students' conceptual understanding, adaptability and professional ethics. AI tools can help map skills, generate practice cases and analyse performance data, but the alignment between enterprise post requirements and curriculum still depends on human negotiation. Teachers are often the people who see where the translation fails: an enterprise standard may

be too narrow for a course, a platform simulation may not match the workshop sequence, or a student assessment may reward completion speed while ignoring diagnostic reasoning.

Boundary work also changes professional development. Enterprise participation in teacher training is useful when it gives teachers access to current equipment, standards, work organisation and typical faults. It is less useful when it becomes ceremonial cooperation with limited curricular influence. Liaoning's response on teacher professional skills explicitly states that leading enterprises and industry resources are invited to participate in training programme design, course teaching and hands-on practice, with the aim of opening production and technical resources to vocational education^[15]. This provides a concrete route for teacher capacity reconstruction, but its effect depends on whether teachers return from enterprise training with time and authority to revise syllabi, teaching materials and assessments.

The same point applies to AI. If enterprises use AI-supported production planning, predictive maintenance, digital twins, intelligent quality inspection or data-based customer management, teachers need to understand these uses as occupational practices rather than abstract examples of digitalisation. School-enterprise co-development should therefore involve joint selection of AI-supported teaching cases, joint validation of digital resources and joint assessment of student competence. Teachers become the hinge of this process. They are neither passive recipients of enterprise requirements nor independent designers detached from work. Their role is to make collaboration educationally meaningful.

3.4 Liaoning-Specific Pathways for Implementation

The fourth finding concerns implementation pathways for Liaoning. The province's context favours a model in which digital teacher development is linked to regional industrial clusters. The education plan already emphasises digital vocational education, new technologies and school-enterprise textbook development^[11]. Later provincial materials describe professional capability maps, revised training plans, curriculum reconstruction, process-based evaluation and capability portraits, with 94 professional maps initiated and new or withdrawn higher vocational majors adjusted according to quality and industrial needs^[19]. These measures point to a data-informed approach to curriculum governance. The challenge is to ensure that teacher development is built into the same governance system rather than treated as a separate training task.

A first pathway is to organise teacher development around professional groups rather than around isolated digital tools. For example, teachers in equipment manufacturing, petrochemicals, robotics, shipbuilding, health services and digital commerce face different AI and data practices. A single AI training course may introduce prompt writing, but it cannot show how AI changes fault diagnosis in smart manufacturing or compliance documentation in a service industry. Professional-group training can link tool use with occupational tasks, enterprise standards and student assessment. It also allows teachers within a specialty to build shared repositories of verified cases, rubrics and simulation scripts.

A second pathway is to establish joint validation routines. Whenever an AI-generated case, simulation, assessment rubric or digital resource is introduced into a vocational course, it should be checked by at least two forms of expertise: pedagogical expertise from teachers and occupational expertise from enterprise partners or industry mentors. The purpose is not to slow down innovation but to prevent inaccurate or inauthentic resources from becoming normalised. This is particularly important for professional courses where students learn procedures, safety norms, technical standards or client-facing judgement. Joint validation also gives enterprises a substantive role in curriculum quality instead of limiting them to internship provision.

A third pathway is to use capability maps as teacher learning tools. Professional capability maps are often designed to improve student training alignment, but they can also reveal what teachers themselves need to learn. If a map shows that a programme requires new competence in data-driven inspection, intelligent equipment operation or digital service design, colleges can identify which teachers are prepared, which need enterprise practice, and which need AI pedagogy support. The teacher development plan then becomes evidence-informed. It can be linked to workload, promotion, teaching innovation projects and team teaching rather than being offered as optional workshops.

A fourth pathway is to build regional communities of practice. Liaoning's regional digital transformation seminars and vocational education digital innovation activities indicate that the province is already creating spaces for exchange around online education, digital empowerment, industry-education integration and talent cultivation^[20]. These activities can be strengthened if they move beyond experience-sharing and develop into sustained communities where colleges compare

teaching cases, evaluate AI-supported resources, discuss failed experiments and co-produce guidelines for ethical AI use. Teachers are more likely to adopt demanding innovations when they can see how colleagues in similar disciplines solve practical problems.

Taken together, these pathways suggest a shift in management attention. The key question is not how many teachers attended AI training, but whether teacher development is connected to professional-group curriculum renewal, enterprise validation, capability maps and regional knowledge sharing. Liaoning has several policy instruments that can support this connection. Their effectiveness depends on coordination at the college level: departments need time for collaborative design, teachers need access to enterprise practice, and leadership needs criteria for recognising boundary work and digital resource validation as legitimate professional labour.

4. Discussion

The findings indicate that vocational teacher roles are being reshaped through a layered process. AI application changes the tools and evidence available for teaching. Industry collaboration changes the source and pace of occupational knowledge. Provincial policy changes the organisational channels through which these demands reach colleges. Teacher role reconstruction happens where these layers meet. This is why the language of replacement is misleading. A teacher who only lectures from fixed materials may indeed be vulnerable to automation. A teacher who designs occupationally authentic learning, checks AI outputs, negotiates with enterprises and supports learner judgement performs work that AI cannot absorb as a simple technical function.

The analysis also clarifies why digital competence should not be treated as an individual attribute detached from institutional context. Empirical studies show that teachers' digital competence and technology use are affected by access, confidence, perceived usefulness, workload and professional profile^[7-9]. In vocational education, these factors are intensified by the need to remain occupationally current. A teacher may be willing to use AI but lack enterprise-validated cases. Another may understand industry practice but be overloaded by teaching and administrative work, leaving little time to redesign digital assessment. Capacity reconstruction therefore requires institutional design: workload allocation, team structures, enterprise access, leadership support, ethical rules and shared resource systems.

For vocational colleges, one managerial implication is to treat AI as part of curriculum governance. AI tools should be reviewed through the same quality lens as textbooks, training equipment and assessment standards. Colleges may establish AI resource review groups at the professional-group level, including teachers, enterprise experts, instructional designers and data protection personnel. Such groups can decide which tools are appropriate, what evidence is required before a generated case enters teaching, how students should disclose AI assistance, and how teachers should document their validation work. This approach avoids both uncritical adoption and excessive prohibition.

This governance view also changes how college leaders should allocate resources. Budgeting for AI-enabled vocational education should include time for teachers to test resources, incentives for cross-departmental course teams, travel or release time for enterprise practice, and support staff who understand data security and instructional design. Without these managerial conditions, teachers may be blamed for slow adoption even when the real constraint is organisational. A college that values only visible digital outputs will encourage quick production of materials; a college that values professional validation will invest in the less visible work that makes those materials credible.

A second implication concerns professional development. Traditional short workshops are unlikely to produce deep change unless they are attached to classroom redesign. International teacher policy increasingly emphasises sustained support, collaboration and working conditions rather than isolated training events^[21,22]. For Liaoning's vocational colleges, this suggests that AI-related training should be organised as design cycles: teachers identify a course problem, study the relevant occupational tasks, work with enterprise partners, build or adapt AI-supported materials, pilot them with students, review evidence and revise. The output is not a certificate of attendance but an improved teaching unit, rubric, simulation or workplace learning task.

A third implication is that ethical AI use should be normalised as part of vocational professionalism. UNESCO's recent discussion of AI and the future of education stresses that AI brings dilemmas around equity, misinformation, agency and

human-machine co-creation^[23]. In vocational education, ethical AI is not an abstract topic. Students may use AI to produce technical reports they do not understand; teachers may use automated scoring without explaining criteria; colleges may collect learner data without clear limits; enterprises may provide proprietary process information that should not be entered into open AI systems. Teachers need institutional rules, but they also need ethical judgement cultivated through cases and peer discussion.

The analysis offers a modest theoretical contribution. It reframes vocational teacher role change as mediated capacity reconstruction. AI application and industry collaboration are not parallel variables acting separately. AI becomes educationally valuable when it is connected to authentic work and teacher judgement; industry collaboration becomes pedagogically valuable when teachers can translate enterprise knowledge into curriculum and assessment. The teacher's reconstructed capacity therefore sits between technology, workplace and learner. This position is demanding because it requires the teacher to hold together different standards of validity: technical correctness, pedagogical accessibility, occupational authenticity and ethical acceptability.

There are also risks. First, teacher capacity reconstruction may become another layer of workload without adequate time or recognition. If teachers are asked to design AI resources, maintain enterprise links, update courses and manage data while their teaching load remains unchanged, reform may produce fatigue rather than innovation. Second, school-enterprise cooperation may privilege large firms and visible technologies, while smaller enterprises and service occupations receive less attention. Third, AI tools may standardise teaching materials in ways that reduce local adaptation. Liaoning's industrial diversity requires plural models, not a single digital template.

Future empirical research should therefore examine how teachers actually experience these reforms. Interviews, classroom observations and design-based studies could test whether the four dimensions identified in this paper appear in practice. Quantitative studies could examine whether participation in enterprise-validated AI curriculum design is associated with teachers' perceived efficacy, student engagement or employment-relevant learning outcomes. Comparative studies across provinces could also identify whether regions with stronger industry-education networks achieve more meaningful AI integration than regions where digital transformation is led mainly by platform procurement.

5. Conclusions

The central problem of AI-enabled vocational education is not whether teachers remain necessary. They do. The more precise problem is what kind of teacher capacity is needed when occupational knowledge, digital tools and school-enterprise collaboration are all changing at once. This study argues that the answer is capacity reconstruction. Vocational teachers in Liaoning and similar industrial regions need to become curriculum translators, AI-supported learning designers, validators of technical and pedagogical quality, brokers of school-enterprise knowledge and stewards of ethical digital practice.

Liaoning's policy environment provides several foundations for this reconstruction: digital vocational education, modern apprenticeship, school-enterprise training bases, industry-education communities, professional capability maps and digital-intelligence literacy initiatives. These foundations should be linked into a coherent teacher development system. The practical priority is to move from tool-centred training to professional-group curriculum renewal; from symbolic enterprise participation to joint validation of learning resources; from platform use statistics to evidence of authentic learning; and from informal AI experimentation to ethical governance.

AI can support vocational education when it extends teachers' capacity to design, diagnose and personalise learning. It weakens vocational education when it is used as a shortcut around professional judgement. The future role of vocational teachers is therefore not smaller but more relational and more accountable. Teachers must work across the borders of classroom, enterprise and platform. For college leaders and policymakers, the task is to make that work visible, supported and institutionally valued.

Funding

No

Conflict of Interests

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

Reference

- [1] UNESCO. (2022). Transforming technical and vocational education and training for successful and just transitions: UNESCO strategy 2022-2029. UNESCO.
- [2] State Council of the People's Republic of China. (2019). National vocational education reform implementation plan. Retrieved from http://www.gov.cn/zhengce/content/2019-02/13/content_5365341.htm
- [3] Standing Committee of the National People's Congress. (2022). Vocational Education Law of the People's Republic of China (revised). Retrieved from https://en.npc.gov.cn.cdurl.cn/2022-04/20/c_909970.htm
- [4] UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.
- [5] UNESCO. (2024). AI competency framework for teachers. UNESCO.
- [6] OECD. (2021). Teachers and leaders in vocational education and training. OECD Publishing. <https://doi.org/10.1787/59d4fbb1-en>
- [7] Xin, Y., Tang, Y., & Mou, X. (2024). An empirical study on the evaluation and influencing factors of digital competence of Chinese teachers for TVET. PLOS ONE, 19(9), e0310187. <https://doi.org/10.1371/journal.pone.0310187>
- [8] Cattaneo, A. A. P., Antonietti, C., & Rauseo, M. (2022). How digitalised are vocational teachers? Assessing digital competence in vocational education and looking at its underlying factors. Computers & Education, 176, 104358. <https://doi.org/10.1016/j.compedu.2021.104358>
- [9] Cattaneo, A. A. P., Antonietti, C., & Rauseo, M. (2025). How do vocational teachers use technology? The role of perceived digital competence and perceived usefulness in technology use across different teaching profiles. Vocations and Learning, 18, 5. <https://doi.org/10.1007/s12186-025-09359-4>
- [10] Roll, M. J. J., & Ifenthaler, D. (2021). Multidisciplinary digital competencies of pre-service vocational teachers. Empirical Research in Vocational Education and Training, 13, 7. <https://doi.org/10.1186/s40461-021-00112-4>
- [11] Liaoning Provincial People's Government. (2022). Liaoning Province 14th Five-Year Education Development Plan. Retrieved from https://www.ln.gov.cn/web/zwgkx/zfwj/szfbgtwj/zfwj2011_148489/A0EA15E046774FD3A6A8118E68BFB7FF/index.shtml
- [12] Liaoning Provincial Department of Education. (2024). Liaoning: Three focus areas empower regional comprehensive revitalization Retrieved from <https://jyt.ln.gov.cn/jyt/gk/zfxxgk/fdzdgknr/zdmsxx/ssqk/2024122416062790022/index.shtml>
- [13] Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- [14] Krippendorff, K. (2018). Content analysis: An introduction to its methodology (4th ed.). SAGE Publications.
- [15] Liaoning Provincial Department of Education. (2024). Reply to Proposal No. 1244 on improving the professional skills of vocational college professional-course teachers. Retrieved from <https://jyt.ln.gov.cn/jyt/gk/zfxxgk/fdzdgknr/jyta/srddb/jy/2024n/2024080915360447952/index.shtml>
- [16] Ministry of Education of the People's Republic of China. (2022). Teacher Digital Literacy (JY/T 0646-2022). Ministry of Education.
- [17] Ministry of Education of the People's Republic of China, National Development and Reform Commission, Ministry of Industry and Information Technology, Ministry of Finance, & National Data Administration. (2026). AI + Education Action Plan. Retrieved from https://www.nda.gov.cn/sjj/zwgk/zcfb/0413/20260410162428109058250_pc.html
- [18] Fejes, A., & Kopsen, S. (2014). Vocational teachers' identity formation through boundary crossing. Journal of Education and Work, 27(3), 265-283. <https://doi.org/10.1080/13639080.2012.742181>
- [19] Liaoning Provincial Department of Education. (2025). Reply to Proposal No. 0654 on constructing a cross-boundary integrated vocational education ecosystem. Retrieved from <https://jyt.ln.gov.cn/jyt/gk/zfxxgk/fdzdgknr/jyta/szxta/2025n/2025090509253022550/index.shtml>

- [20] Liaoning Provincial Department of Education. (2025). Reply to Proposal No. 0620 on improving vocational education teachers' digital-intelligence literacy and accelerating the development of new quality productive forces in Liaoning. Retrieved from <https://jyt.ln.gov.cn/jyt/gk/zfxxgk/fdzdgknr/jyta/szxta/2025n/2025090509244126607/index.shtml>
- [21] OECD. (2024). Education policy outlook 2024: Reshaping teaching into a thriving profession from ABCs to AI. OECD Publishing. <https://doi.org/10.1787/dd5140e4-en>
- [22] OECD. (2025). Results from TALIS 2024: The state of teaching. OECD Publishing. <https://doi.org/10.1787/90df6235-en>
- [23] UNESCO. (2025). AI and the future of education: Disruptions, dilemmas and directions. UNESCO.

Middle Managers and AI-Assisted Performance Appraisal in Chinese Enterprises: Explainability, AI Literacy, and the Case for Human Oversight

Peng Zhang, Rozaini Binti Rosli*

School of Business & Management, Lincoln University College, Petaling Jaya 47301, Selangor, Malaysia

**Corresponding author: Rozaini Binti Rosli, rozaini@lincoln.edu.my*

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: As AI-based appraisal tools move from pilot projects into daily HR routines, middle managers face a practical dilemma. They are asked to use algorithmic scores and model-generated recommendations in rating, feedback, promotion, and compensation decisions, yet they remain accountable for explaining those decisions to employees. This article offers a structured review of research on AI-assisted performance appraisal, algorithmic management, explainability, AI literacy, and procedural fairness. Rather than treating managerial confidence as a simple matter of technological acceptance, the article argues that what matters is calibrated reliance: the ability to use AI outputs seriously without surrendering judgment to them. Three conditions recur across the literature. First, explainability reduces procedural opacity and makes appraisal outcomes easier to defend. Second, AI literacy equips managers to interpret outputs, detect limitations, and resist both blind trust and reflexive rejection. Third, human oversight preserves accountability, contextual correction, and respectful treatment in employee-facing decisions. Building on these themes, the article develops an integrative framework linking explainability, AI literacy, and oversight to middle managers' reliance on AI-assisted appraisal. The discussion then considers practical implications for Chinese enterprises, where performance evaluation often carries strong consequences for pay, promotion, and internal mobility. The paper contributes to the AI-HRM literature by shifting attention from adoption alone to the managerial conditions under which AI-supported appraisal becomes usable, legitimate, and organizationally defensible.

Keywords: AI-Assisted Performance Appraisal; Middle Managers; Explainability; AI Literacy; Human Oversight; Algorithmic Fairness; Chinese Enterprises

Published: Apr 26, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1342>

1. Introduction

Artificial intelligence has moved from the margins of human resource management into core managerial routines. What began with recruitment screening and workforce analytics is now extending into performance appraisal, where AI systems can summarize behavioral traces, detect patterns across large datasets, flag performance deviations, and recommend ratings or follow-up actions^[1]. For organizations, the attraction is obvious: standardization, scale, and speed. For middle managers, the situation is less straightforward. They are often the people who must interpret a score, turn it into feedback, and defend its consequences to both employees and senior leadership.

That position matters. Performance appraisal is not a purely technical exercise. It is also a social and political one. An

appraisal affects pay progression, promotion, developmental opportunities, and informal standing inside the firm. Even when an algorithm appears to improve analytical consistency, the final decision still unfolds in a relationship between manager and employee. This is why AI in managerial work is better understood as a problem of human-AI symbiosis than as simple substitution ^[2]. The practical question is not whether algorithms can process performance data, but under what conditions managers can rely on algorithmic outputs without abandoning accountability.

Recent research has made this tension harder to ignore. Studies of algorithmic management show that algorithms do not merely support work; they can also redistribute control, discretion, and authority in contested ways ^[3]. Parallel reviews in AI and HRM have drawn attention to the promise of AI-enabled decision systems, but they also note persistent problems of opacity, fairness, ethical responsibility, and managerial capability ^[4-6]. These concerns become especially sharp in performance appraisal because the decision is both consequential and visible. Unlike a back-office optimization model, an appraisal score must be explained to a human being who is personally affected by it.

The present article argues that middle managers' use of AI-assisted appraisal is shaped less by abstract enthusiasm for technology than by three linked conditions: explainability, AI literacy, and human oversight. Explainability matters because opaque outputs are difficult to justify, difficult to challenge, and difficult to communicate. AI literacy matters because managers need enough conceptual understanding to read model outputs critically rather than ritualistically. Human oversight matters because appraisal decisions rarely fit the clean boundaries of structured data; they involve context, exception handling, and interpersonal legitimacy.

This paper offers a structured review of recent scholarship and develops a framework of calibrated reliance. Calibrated reliance refers to a mode of use in which managers neither reject AI by default nor defer to it automatically. Instead, they treat AI as a serious but revisable input. The article contributes to the literature in three ways. First, it brings together streams that are often discussed separately: AI in HRM, algorithmic fairness, explainability, AI literacy, and human oversight. Second, it recenters middle managers as the key interpretive layer in AI-assisted appraisal. Third, it translates these insights into implications for Chinese enterprises, where digital HR systems are expanding while appraisal decisions remain tightly tied to managerial authority and employee relations.

2. Review Design and Analytical Lens

This article is a structured review rather than a meta-analysis. It draws mainly on peer-reviewed work published between 2018 and 2025, the period in which AI-assisted HRM, algorithmic management, explainable AI, and AI literacy became a more coherent stream of management and information-systems research. The review was guided by keyword combinations such as 'AI-assisted performance appraisal', 'algorithmic management', 'algorithmic fairness', 'explainability', 'AI literacy', 'human oversight', and 'employee performance evaluation'. Priority was given to studies that speak directly to managerial judgment, employee reactions, organizational design, or the use of AI in people management.

The review deliberately excludes purely technical papers that optimize model performance while saying little about how managers use, communicate, contest, or take responsibility for algorithmic outputs. That choice follows the central premise of this paper: in performance appraisal, the organizational problem is not only predictive accuracy. It is also interpretability, legitimacy, and the translation of machine output into managerial action.

Three analytical questions organize the discussion. First, what makes an AI-supported appraisal understandable enough for a manager to explain and defend? Second, what kind of literacy allows managers to use AI outputs critically rather than passively? Third, what forms of human oversight preserve accountability without reducing oversight to a symbolic gesture? These questions are closely aligned with the literature on opacity in algorithm-based HRM, perceptions of procedural fairness, and the design of human-AI decision systems ^[7,10-12].

Table 1. Analytical themes guiding the review

Theme	Core question	Managerial relevance
Explainability and opacity	Can managers understand, justify, and communicate how the system reached its recommendation?	Without some level of intelligibility, managers struggle to defend ratings and employees are more likely to question the process ^[7,9,11] .

Theme	Core question	Managerial relevance
Fairness and employee reactions	How do employees interpret AI-based evaluation relative to human judgment?	Appraisal acceptance depends on perceived fairness, respectful treatment, and whether the process looks contestable rather than arbitrary ^[10,19,20] .
AI literacy	What do managers need to know in order to use AI outputs well?	Useful literacy is not programming skill alone; it includes recognizing data limits, model assumptions, and appropriate use conditions ^[16-18] .
Human oversight	When should managers follow, revise, or override the system?	Oversight matters when structured indicators miss context, when exceptional cases arise, and when responsibility must remain attributable ^[5,6,8] .
Reliance dynamics	Why do some users distrust algorithms while others over-trust them?	Managerial use sits between algorithm aversion and algorithm appreciation, making calibration more important than simple acceptance ^[12-15] .

3. Explainability and the Problem of Opaque Appraisal

Opacity is one of the defining problems of algorithm-based HRM. Langer and Konig describe opacity not as a purely technical defect but as a multi-stakeholder issue involving users, affected employees, deployers, developers, and regulators ^[7]. In appraisal settings, opacity becomes immediately practical. A middle manager who cannot say why the system produced a score is left with two weak options: either repeat the output as if it were self-validating, or quietly replace it with intuition. Neither option produces robust organizational legitimacy.

Research on algorithmic decisions helps explain why this matters. Lee's work on perceptions of algorithmic decisions shows that algorithmic judgments are often seen as less fair and less trustworthy than human ones, especially when the decision affects people in recognizable social contexts ^[9]. The problem is not only that the system may be wrong. It is also that people have difficulty locating intention, reason, and recourse in an algorithmic decision. In performance appraisal, that difficulty has immediate consequences because employees do not simply receive a score; they interpret what the score says about their standing, effort, and future prospects.

The design literature makes a similar point from the organizational side. Robert and colleagues argue that fair AI in employee management must be designed around distributive, procedural, and interactional justice, rather than around statistical performance alone ^[10]. This is particularly relevant for appraisal. A system may appear distributively efficient by generating consistent scores, yet still fail procedurally if managers cannot explain the pathway from data to outcome, or interactionally if the decision is delivered without recognition of employee context. Explainability, then, should not be understood as a technical add-on. It is part of procedural fairness.

Shin's work strengthens this argument by showing that explainability and causability shape trust, perception, and acceptance ^[11]. In managerial practice, explainability does at least three things. First, it gives managers a basis for checking whether the output is coherent with the case at hand. Second, it helps them translate system logic into a form that can be discussed with employees. Third, it makes disagreement possible. That last point is easy to overlook. A fully opaque system does not merely reduce understanding; it also weakens the conditions for meaningful challenge.

At the same time, explainability should not be romanticized. Not every model can be rendered transparent in a way that is both accurate and managerially useful. And more explanation is not always better if it overwhelms the user with technical detail. For middle managers, what matters is actionable explanation: enough clarity to identify the main drivers of a recommendation, enough traceability to see what data mattered, and enough procedural documentation to justify following or modifying the recommendation. That threshold is lower than full model transparency but higher than mere faith in the vendor or system owner.

Seen this way, explainability is less about turning every manager into a data scientist than about preserving the chain of managerial reasoning. If a manager cannot reconstruct why a recommendation deserves weight, the system may still calculate, but it cannot genuinely support judgment. In a domain as consequential as performance appraisal, that weakness is likely to undermine decision confidence, employee acceptance, or both.

4. AI Literacy as Managerial Interpretive Capacity

Explainability alone does not solve the managerial problem. Even a relatively transparent system can be misread, over-trusted, or rejected if the user lacks the conceptual tools to interpret it. This is where AI literacy enters the discussion. In many organizations, literacy is still treated too narrowly, as if it meant coding expertise or technical familiarity with machine learning pipelines. For appraisal practice, that is the wrong benchmark. The relevant question is whether managers know enough to use the system with discernment.

Behavioral research helps clarify why this matters. Users do not approach algorithms neutrally. Some display algorithm aversion, especially after seeing an algorithm make a mistake^[13]. Others can be induced to rely more on algorithms when they retain even a small degree of control over the final output^[14]. Still other studies document algorithm appreciation, showing that people sometimes prefer algorithmic advice to human advice when they believe the model is competent and objective^[15]. Together, these findings suggest that managerial responses to AI are unstable rather than fixed. Managers may oscillate between skepticism and deference depending on task design, error experience, and perceived agency.

AI literacy provides a way of stabilizing that response. Long and Magerko define AI literacy in terms of the competencies people need to interact with and critically evaluate AI systems^[16]. Ng and colleagues extend the idea by framing AI literacy as a multidimensional construct that includes conceptual understanding, critical interpretation, and awareness of social and ethical implications^[17]. Applied to middle managers, this means literacy should include at least four practical capacities: understanding what kind of data feed the model, recognizing that outputs are probabilistic rather than self-evident truths, identifying when important contextual information is missing, and knowing when escalation or override is warranted.

This matters because appraisal often involves mixed evidence. Some elements of performance are highly structured: sales volume, task completion speed, attendance regularity, response times, or customer ratings. Other elements are not. Mentoring effort, crisis handling, team stabilization, political judgment, or the repair of damaged client relationships often leave weaker digital traces. A manager with low AI literacy may confuse data richness with evaluative completeness. A manager with stronger literacy is more likely to treat the model as informative but partial.

Recent work at the executive level is useful here. Pinski, Hofmann, and Benlian show that AI literacy in top management is related to AI orientation and implementation ability^[18]. The same logic applies one level down. Middle managers do not need to build the system, but they do need enough literacy to use it responsibly. In appraisal settings, literacy is therefore a form of interpretive capacity. It helps managers distinguish signal from measurement convenience, and confidence from unwarranted certainty.

This point also reframes training. Many firms introduce AI tools with short operational briefings focused on dashboards, interfaces, and compliance. That is not enough. Managerial training should include model boundaries, common sources of error, bias risks, and case-based exercises on when to question the system. Without that layer, explainability may exist on paper while interpretive competence remains thin in practice.

In short, AI literacy helps middle managers avoid two symmetrical errors: refusing useful analytical support simply because it is algorithmic, and surrendering judgment simply because the output looks precise. Performance appraisal demands a position between these extremes. That position is learned, not automatic.

5. Why Human Oversight Still Matters

The third condition is human oversight. The phrase is used so widely in AI governance that it risks becoming ceremonial. In practice, oversight only matters if humans have both the authority and the competence to question, revise, or contextualize the output. If the manager is expected to rubber-stamp the score, 'human in the loop' becomes a legitimacy device rather than a real safeguard.

The ethics literature on AI and people management makes this point clearly. Varma and colleagues argue that AI in people management raises enduring questions of responsibility, fairness, and moral legitimacy that cannot be solved by efficiency claims alone^[5]. Rodgers and coauthors similarly emphasize the need to think about ethical decision architectures in HR processes rather than treating AI as a neutral computational layer^[6]. These arguments are directly relevant to appraisal

because appraisal decisions are among the most consequential decisions a manager communicates face to face.

Empirical studies reinforce the point. Tong and colleagues show that AI feedback can improve employee performance, but they also find that how AI is deployed and disclosed matters for employee response^[8]. This suggests that performance gains do not remove the communicative and relational dimension of feedback. A system may optimize recommendation quality while still creating friction if employees feel reduced to a score or monitored without meaningful explanation.

Even more directly, Qin and coauthors find that employees may perceive AI as fairer and more accurate than the average human manager in structured, data-intensive evaluation tasks^[19]. Yet their study also shows why human managers remain central: when employees are evaluated by human managers, perceived fairness plays a first-order role in shaping outcomes. In other words, managers do not retain importance because they are always more accurate. They retain importance because fairness in organizational life is not exhausted by calculative precision.

Chun and colleagues push this further by showing what algorithmic evaluation often fails to provide: respectful treatment and individualized consideration^[20]. That finding is critical for performance appraisal. Employees do not judge an appraisal only by whether the number appears unbiased. They also care whether the decision process acknowledges effort, circumstance, growth trajectory, and the feeling of having been treated as a person rather than a datapoint. This does not mean human managers are inherently fairer. It means that relational legitimacy still matters, and algorithms do not automatically supply it.

For middle managers, then, oversight has at least three substantive functions. First, it is an error-correction mechanism: managers can spot omissions, misclassifications, or unusual cases. Second, it is a contextualization mechanism: they can interpret results in light of organizational realities that are weakly captured in the data. Third, it is an accountability mechanism: they can explain the basis of the decision and bear responsibility for its use. These functions are especially important when appraisal outcomes affect salary, bonus allocation, promotion, or disciplinary action.

A useful oversight model is therefore neither full managerial discretion nor full algorithmic autonomy. It is a disciplined hybrid. Managers should be able to override the system, but overrides should require written reasons. Likewise, following the system should also be documented when the case is consequential. That two-sided documentation matters because it prevents oversight from becoming selective or cosmetic. It also produces an audit trail that can support learning, review, and later system improvement.

6. A Framework of Calibrated Reliance

The literature reviewed above suggests that middle managers' use of AI-assisted appraisal is best understood through the lens of calibrated reliance. Calibrated reliance is not simple trust. It is a practical equilibrium in which the manager assigns weight to AI output in proportion to its intelligibility, task fit, and evidentiary adequacy, while still preserving room for contextual judgment. This concept helps move the discussion beyond the familiar opposition between acceptance and resistance.

The framework proposed here links three inputs to calibrated reliance. The first is explainability. When the system offers intelligible reasons, salient input features, and some trace of how the recommendation was formed, managers are better able to judge whether the output deserves confidence. Explainability also supports communication with employees and creates conditions for contestability^[7,9,11].

The second input is AI literacy. Managers need enough knowledge to understand what the output means, what it does not mean, and where the model's blind spots are likely to be. Literacy therefore moderates the practical value of explainability. A transparent output is of limited use if the manager cannot interpret it, and a complex output is especially risky when users mistake statistical confidence for situational adequacy^[13-18].

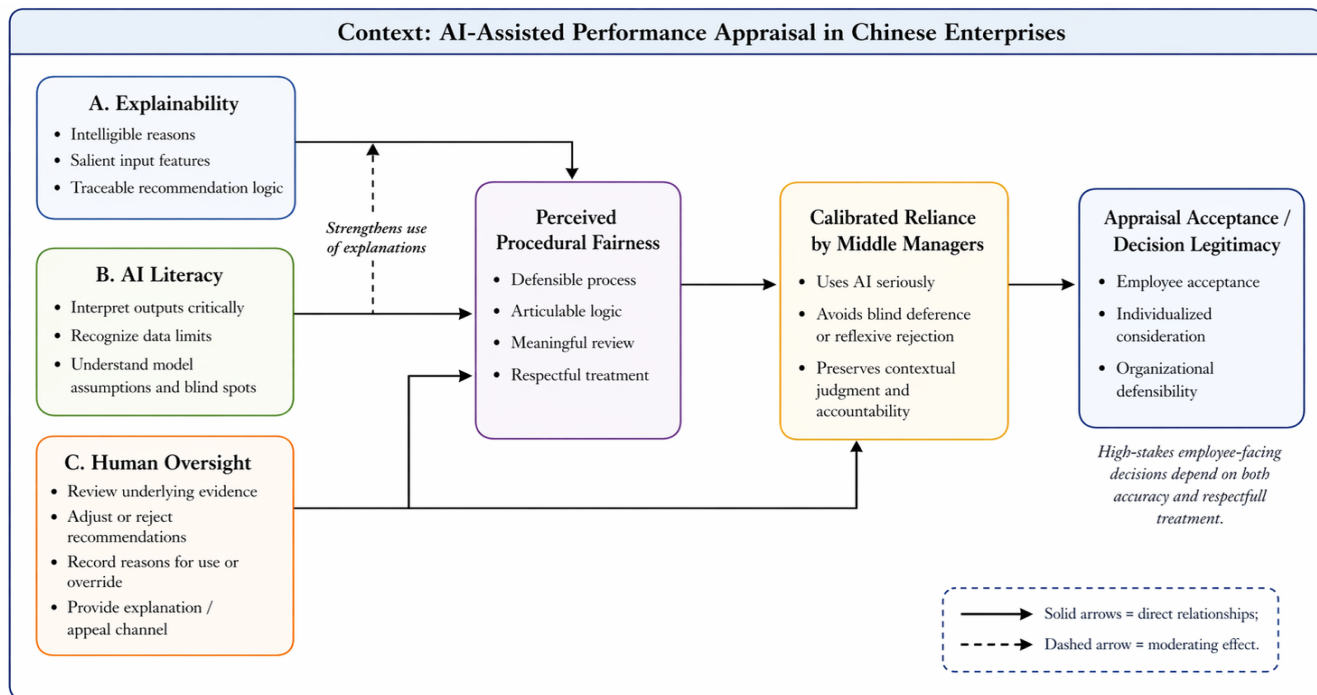
The third input is human oversight. Oversight converts AI from an authority that closes judgment into a tool that supports it. In the framework proposed here, oversight includes the ability to review underlying evidence, adjust or reject recommendations, record reasons, and provide employees with a credible avenue for explanation or appeal. This is what turns managerial involvement into a real governance mechanism rather than a symbolic presence^[5,6,19,20].

These three conditions converge through perceived procedural fairness. Managers are more likely to rely on AI outputs when the process looks defensible, when the logic can be articulated, and when meaningful review remains possible. Employees are more likely to accept AI-supported appraisal when they see that the manager has neither hidden behind the system nor

ignored its analytical value. Calibrated reliance, then, is relational as well as cognitive. It concerns the manager's confidence in using the system and the employee's willingness to treat the decision as legitimate.

On this basis, four propositions can guide future empirical work. First, greater explainability is likely to increase middle managers' reliance on AI-assisted appraisal by improving perceived procedural fairness. Second, AI literacy is likely to strengthen the positive effect of explainability because literate managers can make better use of the information provided by the system. Third, meaningful human oversight is likely to increase calibrated reliance by preserving accountability and contestability. Fourth, in employee-facing appraisal contexts, the effect of AI outputs on decision acceptance is likely to depend not only on perceived accuracy but also on whether the process leaves room for respectful treatment and individualized consideration.

Figure 1. Framework of calibrated reliance in AI-assisted performance appraisal



7. Implications for Chinese Enterprises

The framework has particular relevance for Chinese enterprises, many of which are expanding digital HR infrastructures while also operating in organizational environments where appraisal carries substantial consequences for career progression. In such settings, middle managers sit at a sensitive junction. They are expected to use new systems, demonstrate alignment with data-driven management, and at the same time maintain team stability and interpersonal credibility. That combination makes blind reliance on AI especially risky.

The first implication is design-related. Chinese enterprises should resist the temptation to introduce appraisal systems that produce highly visible scores without adequate explanation layers. A dashboard that displays ranking precision but cannot show the main drivers of the score will place managers in a weak communicative position. For high-stakes decisions, systems should at least display salient indicators, data sources, confidence cues, and a record of whether human adjustment occurred. Managers do not need exhaustive model documentation in daily use, but they do need enough information to defend a recommendation in plain organizational language.

The second implication concerns managerial development. Training should not stop at operational adoption. It should cultivate AI literacy as a managerial competence. Case-based training is especially important: managers should practice reading system outputs, identifying questionable recommendations, and explaining a decision to an employee whose self-assessment differs from the model's conclusion. This kind of training is more demanding than software onboarding, but it is far more relevant to appraisal quality.

The third implication is governance. Firms should formalize two-way accountability around overrides. When managers

depart from AI recommendations, they should state why. When they follow them in consequential cases, they should also record the basis for doing so. This creates a disciplined review process, reduces arbitrary deviations, and produces valuable organizational learning about when the model performs well and when it does not.

The fourth implication is employee-facing procedure. Organizations should offer a clear explanation and appeal channel for AI-supported evaluations. This does not mean every employee should litigate every score. It means there should be an intelligible path for questioning clearly anomalous results or missing contextual evidence. Such channels support legitimacy even when the original recommendation stands.

Finally, firms should be careful about the scope of automation. AI appears most defensible where performance indicators are relatively structured and repeatedly observed. Its role should be more limited where evaluation turns heavily on emergent contribution, tacit coordination, developmental potential, or crisis judgment. A prudent design principle is therefore asymmetry: the more punitive or career-shaping the decision, the stronger the case for deeper human review.

8. Conclusion

AI-assisted performance appraisal is not simply a technical upgrade of an existing HR routine. It changes how performance is seen, measured, explained, and justified inside the organization. For that reason, the key issue is not whether managers are generally ‘for’ or ‘against’ AI. The more useful question is whether organizational conditions allow them to rely on AI outputs in a measured and defensible way.

This review has argued that three conditions are central to that outcome: explainability, AI literacy, and human oversight. Explainability reduces opacity and supports justification. AI literacy enables managers to read outputs critically rather than ritualistically. Human oversight preserves accountability, contextual judgment, and respectful treatment. Taken together, these conditions support calibrated reliance: a mode of use in which AI informs appraisal without replacing managerial responsibility.

The argument matters for research and practice alike. For research, it suggests that future studies should move beyond broad measures of adoption or intention to use and examine how reliance is shaped by explanation quality, literacy levels, and oversight design. For practice, it suggests that organizations should treat appraisal not as a suitable site for full automation, but as a domain in which analytical systems and managerial judgment must be carefully assembled. That lesson is particularly relevant for Chinese enterprises as they deepen digital transformation while trying to preserve fairness, legitimacy, and organizational cohesion.

Funding

No

Conflict of Interests

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

Reference

- [1] Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61(4), 15-42. <https://doi.org/10.1177/0008125619867910>
- [2] Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577-586. <https://doi.org/10.1016/j.bushor.2018.03.007>
- [3] Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410. <https://doi.org/10.5465/annals.2018.0174>
- [4] Pan, Y., & Froese, F. J. (2023). An interdisciplinary review of AI and HRM: Challenges and future directions. *Human Resource Management Review*, 33(1), 100924. <https://doi.org/10.1016/j.hrmr.2022.100924>
- [5] Varma, A., Dawkins, C., & Chaudhuri, K. (2023). Artificial intelligence and people management: A critical assessment through the ethical lens. *Human Resource Management Review*, 33(1), 100923. <https://doi.org/10.1016/j.hrmr.2022.100923>

- [6] Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human Resource Management Review*, 33(1), 100925. <https://doi.org/10.1016/j.hrmr.2022.100925>
- [7] Langer, M., & Konig, C. J. (2023). Introducing a multi-stakeholder perspective on opacity, transparency and strategies to reduce opacity in algorithm-based human resource management. *Human Resource Management Review*, 33(1), 100881. <https://doi.org/10.1016/j.hrmr.2021.100881>
- [8] Tong, S., Jia, N., Luo, X., & Fang, Z. (2021). The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strategic Management Journal*, 42(9), 1600-1631. <https://doi.org/10.1002/smj.3322>
- [9] Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), 1-16. <https://doi.org/10.1177/2053951718756684>
- [10] Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human-Computer Interaction*, 35(5-6), 545-575. <https://doi.org/10.1080/07370024.2020.1735391>
- [11] Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- [12] Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627-660. <https://doi.org/10.5465/annals.2018.0057>
- [13] Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114-126. <https://doi.org/10.1037/xge0000033>
- [14] Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155-1170. <https://doi.org/10.1287/mnsc.2016.2643>
- [15] Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103. <https://doi.org/10.1016/j.obhdp.2018.12.005>
- [16] Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
- [17] Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- [18] Pinski, M., Hofmann, T., & Benlian, A. (2024). AI literacy for the top management: An upper echelons perspective on corporate AI orientation and implementation ability. *Electronic Markets*, 34(1), 1-23. <https://doi.org/10.1007/s12525-024-00707-1>
- [19] Qin, S., Jia, N., Luo, X., Liao, C., & Huang, Z. (2023). Perceived fairness of human managers compared with artificial intelligence in employee performance evaluation. *Journal of Management Information Systems*, 40(4), 1039-1070. <https://doi.org/10.1080/07421222.2023.2267316>
- [20] Chun, J. S., De Cremer, D., Oh, E. J., & Kim, Y. (2024). What algorithmic evaluation fails to deliver: Respectful treatment and individualized consideration. *Scientific Reports*, 14(1), 25996. <https://doi.org/10.1038/s41598-024-76320-1>

Research on the Challenges and Pathways to Breakthroughs in the Development of Consumer Finance in China

Zirong Liu¹, Siyuan Yuan^{2*}

1. Taiyuan Institute of Technology, Taiyuan, Shanxi, 030008, China

2. Chongqing Finance and Economic College, Chongqing, 401320, China

*Corresponding author: Siyuan Yuan, 245316359@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: As a key link connecting the financial system with real-world consumption, consumer finance serves as a vital financial pillar for expanding domestic demand, promoting consumption upgrades and driving high-quality economic development. In recent years, China's consumer finance sector has undergone a profound transformation, shifting from explosive growth to a phase of standardisation. A diverse range of players—including licensed institutions, commercial banks and internet platforms—have collectively built a vast market ecosystem, with digitalisation and scenario-based services becoming increasingly prominent. However, against a complex backdrop of comprehensive regulatory tightening, falling interest rate caps, macroeconomic pressures and intensifying industry competition, China's consumer finance sector faces multiple developmental challenges, including high funding costs, inadequate risk control systems, pressures from multiple borrowing and asset quality, compliance risks associated with loan facilitation models, severe product homogenisation, a lagging credit reporting system and insufficient consumer rights protection. Based on the latest developments in China's consumer finance sector, this paper systematically analyses the industry's deep-seated challenges. By integrating the latest regulatory policies with trends in fintech, it proposes targeted solutions across dimensions including compliant operations, enhanced risk management, deepening engagement with specific scenarios, optimised financing, improved credit reporting, consumer protection and differentiated transformation, with the aim of promoting the sustainable and high-quality development of China's consumer finance industry.

Keywords: Consumer Finance; Development Challenges; Fintech; Risk Management; Pathways to Breakthrough

Published: May 7, 2026

DOI: <https://doi.org/10.62177/apemr.v3i2.1337>

1. Introduction

Against the macroeconomic backdrop of China's transition from an investment-driven to a consumption-driven economy, consumption has been the primary driver of economic growth for several consecutive years. By providing consumer credit services to households, consumer finance effectively alleviates current liquidity constraints and unlocks latent consumer demand; it is a key financial instrument for facilitating the domestic economic cycle, boosting household consumption and improving people's livelihoods. Since China officially launched the pilot scheme for consumer finance companies in 2009, the industry has undergone more than a decade of rapid development. Following 2020, internet-based consumer finance has risen rapidly. By the end of 2025, the number of licensed consumer finance companies nationwide had reached 31, with the

industry's total credit volume exceeding 10 trillion yuan, forming a diversified market structure in which bank credit cards, bank-operated consumer loans, licensed consumer finance companies, and internet platform consumer finance coexist^[1].

In recent years, the regulatory environment for the sector has undergone profound changes. The 2025 revision of the 'Administrative Measures for Consumer Finance Companies', the implementation of new regulations on loan facilitation, the strict cap on the comprehensive annual percentage rate (APR) at 20%, and the full implementation of interest and fee transparency have signalled the sector's departure from extensive growth^[2]. At the same time, increased downward pressure on the macroeconomy has weakened households' debt-repayment capacity, leading to a rise in the sector's non-performing loan ratio and a significant compression of profit margins, marking the onset of a period of transitional growing pains. Drawing upon life-cycle consumption theory and information asymmetry theory, this paper employs literature review, current status analysis and inductive summarisation to examine the industry's current state, dissect its core challenges and propose appropriate pathways, thereby providing practical guidance for the compliant transformation and high-quality development of consumer finance institutions.

2. Current Status and Characteristics of China's Consumer Finance Sector

2.1 Diversification of Market Participants and Continuous Expansion of the Industry

Currently, China's consumer finance market has evolved into a structure where four types of entities—licensed consumer finance companies, commercial banks, internet technology platforms and micro-loan companies—develop in concert. As shown in Tables 1 and Table 2, licensed consumer finance companies, as the mainstay of the industry, numbered 31 by the end of 2025. These are categorised into three major types: bank-affiliated, industry-affiliated and internet-affiliated, with their asset scale and profitability continuing to rise; Leveraging their cost of funds advantage, commercial banks dominate the market through a combination of credit cards and proprietary consumer loans; internet platforms, capitalising on e-commerce scenarios and traffic advantages, have deeply penetrated the long-tail consumer market, significantly enhancing the accessibility of financial services. By the end of 2025, the total credit scale of China's consumer finance industry exceeded 10 trillion yuan, with the breadth and depth of inclusive finance coverage reaching historic highs.

Table 1: The Scale of Development in China's Consumer Finance Industry, 2016–2025

Year	Total consumer credit (trillion yuan)	Number of licensed consumer finance companies (companies)
2016	4.9	18
2017	5.3	19
2018	6.2	23
2019	6.8	24
2020	7.8	27
2021	8.2	28
2022	8.9	30
2023	9.4	30
2024	9.8	31
2025	10.2	31

Source: China Banking and Insurance Regulatory Commission; compiled from annual industry reports

Table 2: Market Share of Key Players in China's Consumer Finance Market in 2025

Market entities	Market share
Commercial banks (credit cards and in-house consumer loans)	62%
Licensed consumer finance company	21%
Consumer finance on internet platforms	13%
Microfinance company	4%

Source: China National Financial Regulatory Administration, 2025 Industry Statistical Bulletin

2.2 Digitalisation and Scenario-Based Lending Emerge as Mainstream Trends

The deep integration of fintech serves as the core driving force behind the industry's development. Technologies such as big data, artificial intelligence and cloud computing have been widely adopted across the entire process—from customer acquisition and risk control to credit approval and debt collection—significantly enhancing the levels of digitalisation, automation and intelligence in business operations ^[3]. At the same time, China's consumer finance sector is gradually moving away from a purely credit-based, scenario-less 'blank lending' model, becoming deeply embedded in real-world consumption scenarios such as e-commerce, tourism and culture, education, home furnishings, automotive and healthcare. In 2024, scenario-based consumer loans grew by 18% year-on-year, with an increasingly evident trend towards a return to the fundamental purpose of serving consumption; the 'finance plus scenario' integration model has effectively reduced the risk of misappropriation of funds ^[4].

2.3 Comprehensive Tightening of Regulation Accelerates the Industry's Transition Towards Standardisation

In recent years, regulatory authorities have continuously introduced policies to standardise industry development and establish a rigorous regulatory framework: firstly, entry barriers have been raised, with the revised *Administrative Measures for Consumer Finance Companies* significantly increasing requirements for registered capital and shareholder qualifications; secondly, business boundaries have been clarified, with new regulations on loan facilitation services introduced to explicitly prohibit the outsourcing of core risk control functions and platform guarantees, thereby severing the chain of risk transmission; Thirdly, strict control over interest rates has been implemented, with interest, service fees and guarantee fees now calculated uniformly to comprehensively reduce hidden high interest rates, and a requirement that the comprehensive annualised interest rate for new loans issued in 2026 does not exceed 20%; fourthly, strict monitoring of capital flows has been enforced to prevent funds from flowing into the property and stock markets in violation of regulations. This series of measures has driven the industry's shift from scale expansion towards compliant, stable and high-quality development.

2.4 The Matthew Effect in the Industry Is Intensifying, with a Marked Polarisation

In this era of competition for existing market share, industry resources continue to concentrate among the leading players. Leading consumer finance companies affiliated with banks have secured a dominant market position thanks to low-cost funding, strong shareholder resources and robust risk management systems, with net profits maintaining steady growth ^[5]; meanwhile, mid-tier and smaller consumer finance institutions, as well as unlicensed platforms, face high funding costs, weak risk management capabilities and difficulties in generating profits. It is predicted that the top five institutions will account for over 70% of the market share in the future, whilst smaller players will face gradual exit from the market, with industry consolidation continuing to accelerate ^[5].

3. Key Challenges Facing the Development of China's Consumer Finance Sector

3.1 Severe Compression of Profit Margins and Structural Imbalances in Funding Costs

The narrowing of profit margins represents the industry's most pressing survival challenge at present, with the core contradiction lying in the divergence of funding costs and the rigid constraints imposed by interest rate caps. Licensed consumer finance institutions primarily rely on interbank borrowing, financial bonds and shareholder deposits for funding; leading institutions, backed by shareholder credit and economies of scale, are able to keep their funding costs within 4%–6%; whereas small and medium-sized institutions, lacking credit backing and stable funding sources, face financing costs as high as 8%–12%. Coupled with fixed operational, customer acquisition and risk provisioning costs, the comprehensive cost of funding for small and medium-sized institutions generally exceeds 16%. Under the 20% interest rate cap, their net interest margin approaches zero or even turns negative, trapping them in a dilemma where 'the larger the scale, the greater the loss'. This structural cost disparity further widens the profitability gap between institutions, posing a severe challenge to the sector's sustainable operations.

3.2 Continued Deterioration in Asset Quality and Concentrated Outbreak of Credit Risk

Macroeconomic volatility and weakening household income expectations have directly led to rising credit risk within the sector. Between 2023 and 2025, the industry's non-performing loan (NPL) portfolio is projected to rise from 14.256

billion yuan to 24.632 billion yuan, with growth rates exceeding 30% for two consecutive years; the pace of risk exposure is significantly outpacing business expansion^[7]. There are three primary root causes of this risk: firstly, widespread multi-lender borrowing, where borrowers' liabilities across platforms accumulate, resulting in actual debt-service ratios far exceeding sustainable levels; secondly, inadequate credit reporting coverage, with a large number of 'long-tail' borrowers lacking credit histories, making precise risk identification difficult^[7]; and thirdly, limited channels for non-performing loan disposal, depressed market prices for asset transfers, and provisioning that continues to erode profits^[7]. The combination of these multiple factors has pushed the industry's asset quality into a downward cycle, with pressure on risk management rising sharply.

3.3 Compliance Breaches in the Loan Facilitation Model and the Erosion of In-House Risk Management Capabilities

Loan facilitation was once the core model driving the industry's rapid expansion, but it also sowed the seeds of compliance and risk management risks. In order to acquire customers quickly, a large number of small and medium-sized institutions outsourced core risk management processes—such as credit approval, risk pricing and post-loan management—to loan facilitation platforms, creating an irregular model characterised by 'platforms directing traffic, institutions providing the capital, and platforms acting as the safety net'. Following the introduction of new regulations prohibiting the outsourcing of core risk control functions and platform guarantees, the shortcomings in institutions' independent risk control capabilities have been laid bare: risk control models reliant on external data, insufficient in-house development capabilities, limited data dimensions, and weak anti-fraud systems have led to a stagnation in business operations and a decline in the quality of new assets. The hollowing out of risk control not only contravenes regulatory requirements but also directly threatens the security of institutions' assets, becoming the greatest obstacle to the industry's transformation.

3.4 Shortcomings in the Credit Reporting System: Data Silos Hinder Risk Control Effectiveness

China's personal credit reporting system centres on the People's Bank of China's credit reporting, yet it suffers from three major shortcomings: insufficient coverage, poor data sharing, and a lack of diversity in data dimensions. On the one hand, long-tail groups such as blue-collar workers, new urban residents and self-employed individuals lack credit records, creating 'blind spots' in risk control; on the other hand, data remains isolated between licensed consumer finance institutions, internet platforms and micro-loan companies, making it impossible to verify users' total liabilities and multiple borrowing in real time, thus creating data silos^[7]. Incomplete and non-interoperable credit information directly leads to low accuracy in institutional risk control, as well as high incidence of adverse selection and moral hazard, constituting the underlying cause of the intractability of credit risk^[7].

3.5 Product and Scenario Homogenisation, Lack of Differentiated Competitive Capabilities

The industry is generally plagued by product homogenisation, superficial application scenarios, and rudimentary service provision. Products are predominantly cash loans and small-value general instalment plans, with highly similar credit limits, terms and interest rates, lacking customised designs for specific customer segments and vertical scenarios; Scenario partnerships are mostly limited to superficial payment instalments, failing to engage deeply in merchant operations, user engagement and risk management cycles, resulting in high risks of misappropriation of funds and low user retention^[4]. Institutions lack core competitiveness and are thus trapped in price wars and a race for traffic, making it difficult to establish long-term competitive advantages or to genuinely serve authentic consumer needs.

4. Pathways to High-Quality Development in China's Consumer Finance Sector

4.1 Licensed Consumer Finance Companies: Optimising Financing Structures and Reducing Overall Funding Costs

Licensed consumer finance companies should make cost reduction and liability optimisation their core strategies. Firstly, they should broaden their financing channels by actively issuing financial bonds, ABS and Tier 2 capital bonds, monetising existing credit assets, and seeking support from the central bank through re-lending and rediscounting to replace high-cost interbank liabilities; Secondly, they should optimise the maturity structure of their liabilities, increasing the proportion of long-term, stable funding to alleviate the duration mismatch arising from borrowing short-term to lend long-term; thirdly,

they should strengthen synergies with shareholder resources, encouraging high-quality shareholders to inject capital and thereby reducing reliance on high-interest market-based financing. By adopting a multi-pronged approach, they can bring overall financing costs under control and drive them down, thereby creating room for profit.

4.2 Licensed Consumer Finance Companies: Rebuilding Autonomous Risk Control and Fortifying Credit Risk Defences

Licensed consumer finance companies must reclaim control over risk management and address technological shortcomings. Firstly, fully implement the new regulations on loan facilitation by independently establishing an end-to-end risk control system covering customer acquisition, credit approval, pricing and post-loan management, whilst eliminating outsourcing of core processes; secondly, utilise big data, artificial intelligence and large-scale models to construct multi-dimensional user profiles, thereby enhancing the ability to identify risks associated with multiple borrowing, excessive indebtedness and fraud; thirdly, establish a rigid liability control mechanism, setting thresholds such as the monthly debt-to-income ratio and total debt ceiling to curb excessive borrowing at source ^[6]. By leveraging technology to empower independent risk control, achieve steady growth whilst maintaining risk under control.

4.3 Commercial Banks: Leveraging Funding Advantages to Deepen Engagement in Scenarios and the Inclusive Finance Market

Commercial banks should adopt a differentiated strategy by capitalising on their three key strengths: low-cost funding, physical branches and a robust customer base. Firstly, they should focus on real-life scenarios such as home furnishings, motoring, cultural tourism, healthcare and green consumption, developing closed-loop instalment products to mitigate the risk of misappropriation of funds ^[4]; Secondly, they should expand into county-level and rural markets, developing dedicated inclusive financial products for new urban residents, blue-collar workers and farmers to enhance service coverage ; thirdly, they should collaborate with offline merchants and industry partners, shifting from merely providing credit to comprehensive ‘finance + scenarios + services’ operations, thereby strengthening user retention and asset quality .

4.4 Regulatory Framework: Improving the Credit Reporting System and Strengthening Compliance and Institutional Safeguards

Regulators are focusing on breaking down data barriers, improving institutional rules and regulations, and standardising market order. Firstly, accelerate the achievement of comprehensive credit reporting coverage, promoting complementarity between the People’s Bank of China’s credit reporting system and market-based credit reporting agencies, and fully incorporating long-tail customer groups and small-scale credit records into credit reports ; secondly, utilise technologies such as privacy-preserving computing to achieve compliant data sharing, break down data silos between institutions, and accurately monitor multiple borrowing ; Thirdly, expediting specialised legislation for consumer finance to elevate the legal standing of regulatory provisions, strictly enforcing licensed operations, clearing out unlicensed and non-compliant institutions, and maintaining a fair competitive environment; fourthly, refining consumer rights protection rules to standardise interest rates, fees, marketing and debt collection practices, and mitigate reputational risks within the industry ^[8].

4.5 Industry-Wide Collaboration: Promoting Product Innovation and Returning to the Essence of Consumer Services

The entire industry should pursue a transformation characterised by differentiation, scenario-based approaches and financial inclusion. Firstly, institutions should position themselves within specific market segments based on their own resources, avoiding homogenisation and internal competition, and developing customised products for young people, family households and the elderly ; secondly, deepen scenario integration, shifting from ‘credit supply’ to ‘consumption empowerment’, and designing products and services around genuine consumer needs ^[8]; thirdly, strengthen financial literacy, enhance users’ awareness of debt and risk, and drive the industry’s transition from a ‘scale-oriented’ approach to a ‘high-quality development-oriented’ approach.

Conclusion

China’s consumer finance sector serves as a core financial pillar for expanding domestic demand, driving economic structural transformation, and boosting household consumption. Following more than a decade of rapid, unsustainable expansion, the

sector has now fully entered a new phase of high-quality development characterised by stringent regulation and structural transformation ^[9]. Currently, the industry's total credit volume has exceeded 10 trillion yuan; however, it simultaneously faces core challenges such as squeezed profit margins, pressure on asset quality, hollowed-out risk control in loan facilitation, siloed credit data, and homogenised product competition. The Matthew effect within the industry continues to intensify, with smaller institutions facing pressure to exit the market.

The root causes of the industry's current difficulties stem from a combination of factors, including reliance on traditional, extensive growth models, a comprehensive tightening of regulatory policies, changes in the macroeconomic environment and household debt landscape, and insufficient integration of fintech with risk management systems. To break through these developmental bottlenecks, licensed consumer finance companies, commercial banks and regulators must collaborate through clear division of labour and targeted efforts: licensed consumer finance institutions should focus on three core areas—cost reduction, risk management and innovation; commercial banks should leverage their capital and scenario advantages to deepen their engagement with inclusive finance and real-economy consumption; and regulators should improve credit reporting, legislation and compliance oversight to foster a healthy market ecosystem.

The industry must remain true to its fundamental purpose of serving household consumption, adhering to four key directions: compliant and prudent operations, technology-enabled risk control, deep integration with real-world scenarios, and sustainable inclusive development. By comprehensively addressing development bottlenecks, the industry can ultimately achieve a fundamental transformation from high-speed, extensive growth to high-quality, sustainable development, thereby continuing to support the implementation of the strategy to expand domestic demand and the stable 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] National Financial Supervisory Administration. (2025). Administrative measures for consumer finance companies (2025 revision).
- [2] National Financial Supervisory Administration. (2025). Notice on strengthening the management of commercial banks' online loan facilitation services and improving the quality and efficiency of financial services.
- [3] Liu, D. (2021). New opportunities for the digital development of consumer finance. *China Finance*, (11), 39–40.
- [4] Lu, R. (2022). A study on the scenario development of consumer finance companies in China. *Cooperative Economy and Technology*, (04), 86–88.
- [5] Yuan, J., & Zhang, L. N. (2021). The development of bank-affiliated consumer finance companies. *China Finance*, (24), 59–60.
- [6] Cheng, X. J., & Fan, Y. P. (2022). Development challenges and solutions for the deep integration of fintech into the consumer finance industry. *Lanzhou Journal*, (11), 34–49.
- [7] Li, Y. Q. (2014). Constraints on the development of China's consumer finance and policy options. *Shandong Social Sciences*, (03), 149–153.
- [8] Zhang, J. (2015). Prospects and strategic choices for the development of consumer finance in China. *Economic Review*, (07), 109–112.
- [9] Huang, D. Z. (2022). Review and outlook of the consumer finance industry. *Financial Panorama*, (12), 22–26.

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.

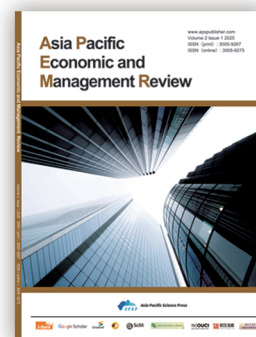
Should you wish to submit your work or express interest in joining our team, please do not hesitate to contact us. You can submit your manuscript or personal profile to info@apspublisher.com or visit our website at www.apspublisher.com for more information.

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.

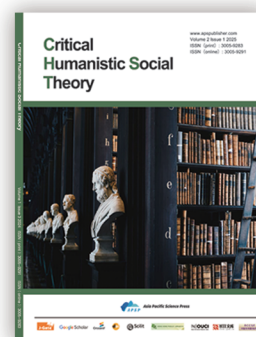
Warm regards
Asia Pacific Science Press

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.



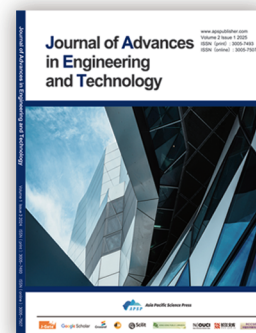
Critical Humanistic Social Theory is an journal that publishes papers specifically using quantitative or qualitative research methods for social science research. The journal encourages scholars to conduct social science theory research from the perspective of social critical theory and emphasizes research 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.

