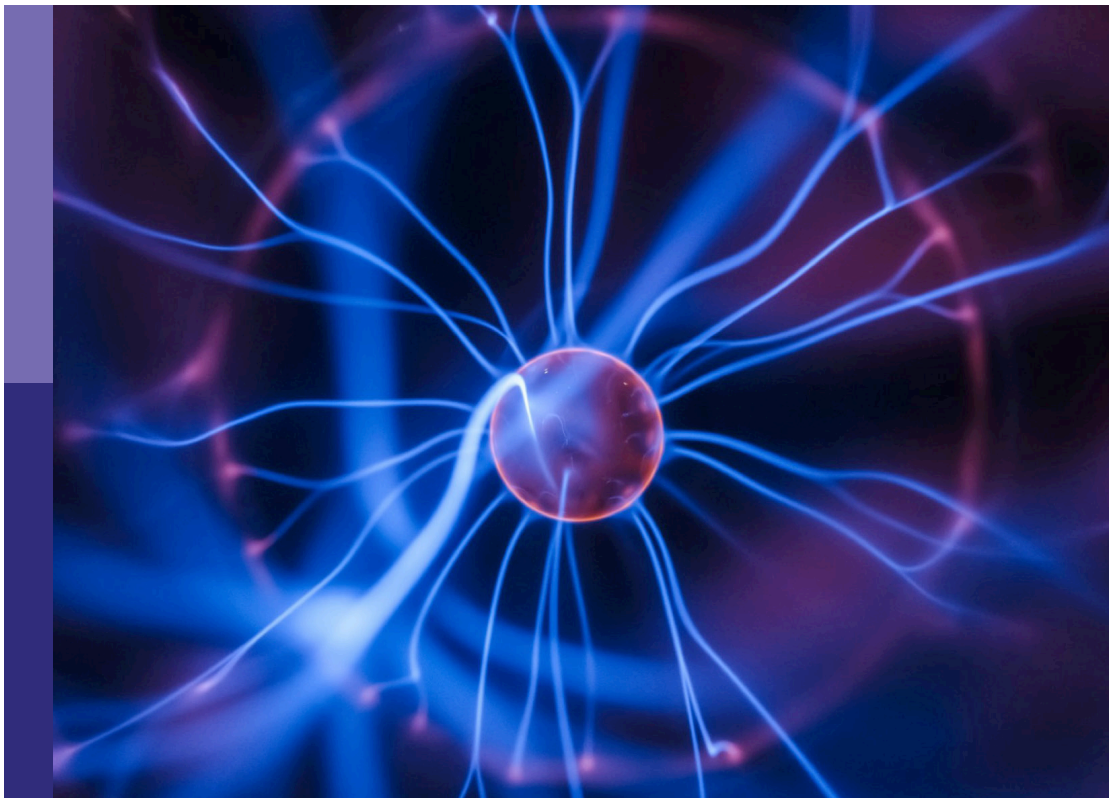
frontiers | Research Topics

Finance and production complex systems

Edited by
Ze Wang, Shupei Huang and Brian Lucey

Published in
Frontiers in Physics



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-7195-8
DOI 10.3389/978-2-8325-7195-8

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Finance and production complex systems

Topic editors

Ze Wang — Capital Normal University, China

Shupei Huang — China University of Geosciences, China

Brian Lucey — Trinity College Dublin, Ireland

Citation

Wang, Z., Huang, S., Lucey, B., eds. (2025). *Finance and production complex systems*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-7195-8

Table of contents

06 Resilience measurement and dynamic evolution of China's provincial financial infrastructure
Guijun Li, Zhongwei Xiong, Chenhuan Kou and Shanshan Li

19 Simulation analysis on the evolution driven model of global copper ore trade network
Qishen Chen, Kun Wang, Yanfei Zhang, Qing Guan, Jiayun Xing, Tao Long, Guodong Zheng, Qiang Li, Zhenqing Li, Xin Ren, Chenghong Shang and Yueran Duan

33 Co-movement forecasting between consumer sentiment and stock price in e-commerce platforms using complex network and entropy optimization
Mingyue Wang, Rui Kong, Jianfu Luo and Wenjing Hao

49 Financial crisis impact on the efficiency evaluation of university research achievements transformation: a study based on fixed and interactive network data envelopment analysis methods
Yuping Ma and Mingqian Xu

63 Systemically important company identification based on topological characteristics of the common share pledge network
Yuyin Ma, Qiong Sun and Chao Wang

77 Evolutionary characteristics and community detection of the international metal minerals trade
Yanfei Zhang, Qishen Chen, Kun Wang, Qing Guan, Jiayun Xing, Tao Long, Guodong Zheng, Mei Li, Chenghong Shang and Jingwen Mao

91 Research on characterization and prediction of bond risk factors based on machine learning: evidence from the China
Yaowen Zhang and Wei Cui

110 The impact of climate risk on the soundness of local commercial banks in China—based on physical risk and transition risk
Yang Xu, Yiqiu Wang, Meng Xie and Chong Guo

130 Research on bank financing strategy of e-commerce supply chain driven by blockchain
Lei Yang and Qiaoming Hou

153 Optimization of bank investment portfolio and debt structure under sustainable finance policies based on mathematical modeling analysis
Guodong Zhang

171 Non-financial enterprises' shadow lending and SME survival: the role of spatial spillovers in regional financial stability
Chen Wang, Wei Zhang, Yongqiang Meng and Xiong Xiong

192 Global geopolitical risk and industrial tail risk in the Chinese stock market: a quantile-on-quantile connectedness approach
Jing Li, Yunzhong Li, Fujun Lai and An Li

210 Study on risk contagion between rare earth and new energy-related metals markets from the perspective of co-production and co-consumption
Renrui Zheng, Hao Nong, Qishen Chen, Liuguo Shao, Hongtao Zhang and Yanfei Zhang

236 Spillover effects in stock index returns within Chinese oil industry chain
Likun Lei, Yu Wang, Liang Wu and Yufeng Mao

249 Does climate risk undermine urban economic resilience? —Evidence from Chinese cities
Di Wu, Qiguang An and Lin Zheng

267 Complex impacts between economic policy and listed oil companies: a dynamics-based approach
Yanzhi Duan, Xingsheng Jing, Haitang Xu, Zhong Wang, Xin Zhang and Yi Wang

279 Unintended consequences of quote limits policy on stock market dynamics: evidence from the MF-X-DMA analysis
Dayong Lv, Yaping Zhou, Jiarui Zhang and Linsen Yin

292 Studies on innovative energy startups' topological roles and their correlation with success: based on temporal networks
Ying Wang and Qing Guan

308 Market state transitions and crash early warning in the Chinese stock market
Wu-Yue Pang and Li Lin

323 Research on the risk contagion of banks holding subordinated debt from the perspective of risk preferences
Shuting Chen, Shanshan Jiang, Yikang Zhuang and Lingyi Meng

335 Blockchain-driven logistics provider guaranteed financing strategy considering risk aversion
Lei Yang and Qiaoming Hou

359 Structural robustness of China's UAV enterprise network: a national and provincial analysis based on business homogeneity
Xuemei Zhang, Jianfeng Guo and Siyao Liu

376 Benefit analysis of regional energy interconnection using system dynamics method targeting carbon neutrality
Ming Lv, Zijing Gao, Han Jiang, Yi Gao and Yuan Gao

395 Real and simulated shocks to the food system: a nested food-economic system model for assessing resilience using a networked input–output approach
Catherine Brinkley, John Ulimwengu and Subhashni Raj

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 19 February 2025
DOI 10.3389/fphy.2025.1556285

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Daniela Feschiyan,
University of National and World Economy, Bulgaria
Raqif Qasimov,
Azerbaijan State University of Economics, Azerbaijan
Xinya Wang,
University of International Business and Economics, China
Xiaofen Tan,
Beihang University, China

*CORRESPONDENCE
Chenhuan Kou,
kouchch@hueb.edu.cn

RECEIVED 06 January 2025
ACCEPTED 29 January 2025
PUBLISHED 19 February 2025

CITATION
Li G, Xiong Z, Kou C and Li S (2025) Resilience measurement and dynamic evolution of China's provincial financial infrastructure. *Front. Phys.* 13:1556285. doi: 10.3389/fphy.2025.1556285

COPYRIGHT
© 2025 Li, Xiong, Kou and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Resilience measurement and dynamic evolution of China's provincial financial infrastructure

Guijun Li[1,2], Zhongwei Xiong[3], Chenhuan Kou[4]* and Shanshan Li[5]

[1]Hebei Normal University, Shijiazhuang, China, [2]Center for Global Economy and Sustainable Development, Central University of Finance and Economics, Beijing, China, [3]School of Management Science and Engineering, Central University of Finance and Economics, Beijing, China, [4]School of Management Science and Information Engineering, Hebei University of Economics and Business, Shijiazhuang, China, [5]Institute of Finance and Economics, Central University of Finance and Economics, Beijing, China

Financial infrastructure provides a fundamental guarantee for the sound and efficient operation of financial markets. This paper constructs a provincial-level financial infrastructure resilience evaluation index system, measures the level of financial infrastructure resilience of 31 provinces in China from 2005 to 2020, reveals the dynamic evolution characteristics of financial infrastructure resilience in the spatial and temporal dimensions, and analyzes the barrier factors affecting the dynamic change of financial infrastructure resilience level. The results of the study show that: Provincial financial infrastructure resilience level overall upward trend, in which resilience of risk and control is in a stagnant growth state, resilience of recovery and stability, resilience of transformation and development is on the rise; The level of financial infrastructure resilience and the rate of change generally show a spatial distribution trend of "low west and high east, north-south concave," with the "wooden barrel effect" evident in the eastern region; Provincial financial infrastructure resilience is in the medium-low coupling stage, affected by the spillover effect is prone to form a "club convergence" phenomenon; The industrial structure, shadow banking risk, real estate liability risk, and fintech level are the main barrier factors to the resilience of financial infrastructure, resilience of risk and control, resilience of recovery and stability and resilience of transformation and development's barrier degree respectively show a rising, declining, stable trend of change. Therefore, the follow-up should promote the coordinated development of regional finance and form a regional financial development pattern with staggered development and distinctive features; prevent the risks of regional financial infrastructure; and enhance the resilience of risk and control and the resilience of transformation and development.

KEYWORDS

financial infrastructure, resilience, dynamic evolution, barrier factors, coupling coordination

1 Introduction

The Chinese government attaches great importance to financial security, and has elevated financial infrastructure to a higher strategic position. In 2015, the

State Council issued the Circular on the Plan for Promoting the Development of Inclusive Finance (2016–2020), which states that "financial infrastructure is an important pillar and platform for improving the operational efficiency and service quality of financial institutions". In September 2019, the 10th Meeting of the Central Committee for Comprehensively Deepening Reform considered and adopted the Work Program for Coordinating and Supervising Financial Infrastructure, which pointed out that financial infrastructure is a fundamental guarantee for the sound and efficient operation of the financial market, and an important hand for implementing macro-prudential management and strengthening risk prevention and control. In March 2020, the People's Bank of China and other six departments jointly issued the Work Program for Coordinating and Supervising Financial Infrastructure The Program further clarified that the scope of coordinated supervision of financial infrastructure includes six types of facilities and their operating institutions, including financial asset registration and custody systems, clearing and settlement systems, trading facilities, transaction report repositories, important payment systems and basic credit systems.

Financial infrastructure undertakes key functions such as payment, clearing, settlement and record keeping in financial market [1], and occupies a pivotal position in the operation of financial market [2, 3]. At present, the research on financial infrastructure at home and abroad is in the stage of transition from the embryonic stage to the development stage. The academic research on financial infrastructure covers the concept and function, risk management, governance and regulation. Risks existing in financial infrastructure have gradually been paid attention to by scholars, including liquidity and operational risks, and facility-associated risks. For example, Berndsen (2017) [4] constructs a financial infrastructure operational risk indicator based on daily transaction data, and then suggests the size of the risk by comparing it with the benchmark number. León (2018) [5] constructs a multilayered network to examine the correlation of systemic risk of Colombia's financial infrastructures, and finds that the core financial institutions are cross-overlapping in the financial network, and their systemic importance is much higher than that of a single value measured in the network. Wendt (2015) [6] argues that central counterparties are highly interconnected with financial institutions and markets and suggests network analysis to assess potential losses and adopt alternative policy measures to reduce systemic risk. The construction of a multi-level financial infrastructure system by strengthening the integrated regulation of financial infrastructure can help prevent and mitigate financial risks [3, 7]. It has been shown that financial infrastructure has an important impact on economic development and financial stability, and in the context of continuous innovation and change in financial markets, risk management of financial infrastructure has become a hot topic.

With the continuous innovation and globalization of financial markets, the traditional risk management model is no longer able to cope with the challenges faced by increasingly complex financial infrastructures. Scholars have gradually explored how to enhance the resilience and recovery capacity of financial infrastructure when it encounters systemic risk from the perspective of resilience [8–10]. For example, Kass-Hanna (2022) [11] argues that improving digital literacy can increase financial infrastructure resilience, and Saligna (2019) [12] proposes to measure financial infrastructure resilience at the individual level. There are relatively few studies on financial infrastructure resilience, and there is still a lack of in-depth analysis and empirical data to support the performance of financial infrastructures in the face of shocks such as systemic risks, natural disasters, and technological failures, as well as their resilience in the aftermath of shock.

Financial infrastructure resilience refers to the ability of financial infrastructure to adjust its own functional structure in a timely manner according to the objective environment in response to the system's own and external risk perturbations, and thus continue to provide critical financial services to the market [1]. Exploring the level of financial infrastructure resilience, dynamic evolution characteristics and resilience obstacle factors can not only enrich the knowledge system and theoretical framework in this field, and deeply understand its operation mechanism, risk management, and impact on the financial market, but also provide a powerful guidance for policy making and facility operation. Based on this, this paper intends to construct a financial infrastructure resilience evaluation index system to measure the level of financial infrastructure resilience in 31 provinces and regions of China from 2005 to 2020, and adopt the Dagum Gini coefficient, the degree of coupling coordination, and the spatial Markov chain model to reveal the dynamic evolution characteristics of the resilience of financial infrastructure. On this basis, the barrier degree model is introduced to analyze and determine the barrier factors affecting the dynamic change of the resilience level. The findings of this paper are conducive to the policy and behavioral adjustment of financial infrastructure, and effectively and efficiently improve the level of financial infrastructure resilience.

2 Methods and data sources

2.1 Methods

2.1.1 Entropy method

This paper adopts the entropy method to measure the weights of the indicators of financial infrastructure resilience level. At the same time, in order to ensure that all the indicators involved in the evaluation are comparable, and to avoid meaningless logarithmic calculation when seeking entropy value, this paper adopts the method of standardization of polar deviation, non-negativity of positive and negative indicators, and uniformly assigns a value of 0.01. The specific methods are as follows:

Positive indicators:

$$X_{ij}^{*} = \frac{X_{ij} - min\left(X_{ij}\right)}{max\left(X_{ij}\right) - min\left(X_{ij}\right)} + 0.01$$

negative indicators:

$$X_{ij}^{*} = \frac{max\left(X_{ij}\right) - X_{ij}}{max\left(X_{ij}\right) - min\left(X_{ij}\right)} + 0.01$$

2.1.2 Dagum Gini coefficient

The Dagum Gini coefficient can decompose the overall differences into intra-region differences, inter-region differences and hypervariance densities, which in turn can reveal the relative

differences due to the sample spatial overlap problem [13–15]. In this paper, the Dagum Gini coefficient and its decomposition method are used to examine the relative differences in the resilience level of China's provincial financial infrastructures, and the relevant calculation formula refers to Zhou et al [16, 17].

$$G = G_w + G_{nb} + G_t$$

$$= \sum_{j=1}^{k} G_{jj} p_j s_j + \sum_{j=1}^{k} \sum_{h \neq j} G_{jh} p_j s_h D_{jh} + \sum_{j=1}^{k} \sum_{h \neq j} G_{jh} p_j s_h \left(1 - D_{jh}\right)$$

2.1.3 Degree of coupling coordination

In this paper, the degree of coupling coordination is used to examine the degree of harmonious development of internal components of financial infrastructure resilience and the dynamic development process of the system from disorder to order and from low level to high level gradually under the joint action of multiple internal and external factors. Define the coupling degree formula as:

$$C = 3 * \frac{1}{3} \sqrt{\frac{A_1 * A_2 * A_3}{\left(A_1 + A_2 + A_3\right)^3}}$$

the coordination degree formula:

$$T = A_1 + A_2 + A_3$$

Degree of coupling coordination formula:

$$D = C * T$$

2.1.4 Spatial Markov chain

This paper introduces the spatial Markov chain [18–20], which classifies the resilience level into four types (low, lower, medium-high, and high) according to the quartiles, denoted by $K = 1, 2, 3, 4$, respectively, and then examines the dynamic evolution of the resilience level of financial infrastructure. In order to better portray the interactions among provinces this paper adopts the spatial geo-economic weight matrix W, so that the financial infrastructure resilience of neighboring provinces $\sum j w_{ij}{}^* y_j$, where w_{ij} is the element in W and y_j is the level value of province j. When considering K spatial lag types as transfer conditions, this will produce $K \times K$ conditional transfer arrays. The meaning is the probability that the resilience level is transferred from type i to type j under the condition that the spatial lag resilience level is of type ω.

2.1.5 Barrier degree model

This paper introduces the barrier degree model, which can effectively measure the degree of influence of each barrier factor on the resilience of financial infrastructure. Where the factor contribution degree u_i indicates the degree of contribution of individual indicators to the overall goal, which is expressed by weights in this paper; the indicator deviation degree Z_{ij} indicates the difference between a single indicator and the optimal value; the barrier degree F_{ij} indicates the degree of influence of a single indicator on the level of financial infrastructure in the provincial area, and the criterion level barrier degree is the sum of barrier degrees of the individual indicators it contains.

$$u_i = w_j$$

$$Z_{ij} = 1 - X_{ij}^*$$

$$F_{ij} = \frac{u_i Z_{ij}}{\sum_i u_i Z_{ij}}$$

2.2 Data sources

This paper explores the spatial-temporal evolutionary pattern of financial infrastructure resilience using 31 provinces in China as the study unit. The data are mainly obtained from official sources such as China Statistical Yearbook, China Financial Yearbook, China Securities Industry Yearbook, while web crawler technology is utilized to obtain supplementary parts of the data. For the missing part of the data, linear interpolation and other methods are used to complete the data. Although most of the data are relatively complete, there are still a small number of missing, such as local government debt data for individual years are missing, this paper uses linear interpolation and other methods to ensure the completeness of the overall data and the accuracy of the analysis.

By analyzing the historical process of China's financial system reform, it can be found that, after the financial system reform in the 1980s and 1990s, the rapid development of financial subsectors and the initial construction of the pattern of sectoral supervision in the 21st century have gradually formed a complete and well-regulated financial infrastructure system. However, with the rapid development of financial integrated management and Internet finance, the layout of financial infrastructure does not match with the rapidly developing financial environment, and at the same time leads to new regulatory challenges. Based on this, this paper selects 2005–2020 as the research cycle to analyze the spatial-temporal evolution pattern of China's financial infrastructure resilience.

3 Construction of the indicator system

3.1 Resilience dimension setting

This paper draws on the theory of economic resilience and sets up financial infrastructure resilience sub-dimensions in combination with financial infrastructure functions, including resilience of risk and control, resilience of recovery and stability, and resilience of transformation and development. There is a high degree of correlation between the operation of the economy and the financial system [21–23]. When the financial system is hit, the service capacity of financial institutions is impaired, leading to a decline in the efficiency of social capital allocation, which in turn affects the operation of the real economy. The theory of economic elasticity originated from the study of the ability of the economic system to cope with external shocks, focusing on the system's ability to resist, recover, and the potential to achieve new development paths through structural optimization and technological innovation [24–26]. Based on the high correlation between the economic system and the financial system, economic resilience provides a feasible perspective for the sub-dimension setting of financial infrastructure resilience. Specifically, as the core pillar of the financial system, the resilience of the financial infrastructure is not only related to

the ability to effectively prevent systemic risks, but also directly affects the recovery speed and transformation potential of the financial system in a crisis [27]. Economic resilience theory provides a multi-dimensional and dynamic evolutionary perspective, from the perspective of economic resilience's shock resistance, recovery and innovation ability, the three dimensions of resilience of risk and control, resilience of recovery and stability, and resilience of transformation and development effectively delineate the resilience of the financial infrastructure, which effectively depicts its adaptive capacity and change potential in the face of complex environments. This framework helps to understand the key role of financial infrastructure in coping with shocks, supporting recovery and promoting transformation, and also provides theoretical support for the stability and development of the financial system in the face of dynamic change.

3.2 Resilience influencing factors

In this paper, with the help of Nvivo14 software, the three-tiered coding system of open coding, axial coding, and selective coding from grounded theory is used to categorize and generalize the factors affecting the resilience of financial infrastructures [28–30]. Open coding is the process of breaking up and then initially aggregating the data information. Based on the research topic, phrases or keywords that may be needed are identified as preliminary conceptual presentations, and then the abstracted concepts are clustered and integrated in constant comparison and form more abstracted categories such as categories. Twenty-six initial categories were obtained through open coding, and due to space constraints, only some of the open coding processes are presented in Table 1.

Axial coding refers to the further work of aggregating the disassembled data material to find new connections that reflect the meanings presented by the categories in a more generalized way. On the basis of open coding, further sorting and generalization is carried out to refine a total of five main categories: financial stability, financial liberalization, macroeconomics, government intervention, and actors. Finally, based on the data availability and word frequency coverage of the initial categories, the system of factors affecting the resilience of financial infrastructure containing 10 initial categories is retained, of which financial stability includes shadow banking scale, insurance depth, and non-performing loan ratio, financial liberalization includes the level of inclusive finance, and fintech level, macroeconomics includes real estate liability risk, and industrial structure, and government intervention includes government debt risk, financial regulatory intensity, and the actors include only information asymmetry.

3.3 Indicator system

Measurement indicators may have impact, resistance, adjustment, and transformation impacts on financial infrastructure, according to the qualitative viewpoints on measurement indicators in the existing literature in the grounded theory, the measurement

TABLE 1 Summary of factors affecting the resilience of financial infrastructure.

Initial scope	Perspectives (simplified)
Insurance premium depth	The greater the depth of premiums, the greater the ability of the financial system to adapt and adjust autonomously
	Insurance companies operate core insurance business without systemic risk
Size of shadow banking	The "shadow banking" system increases the likelihood of risk outbreaks and contributes to the contagion of systemic risk
	Systemic Financial Risks Currently Facing China: Real Estate Risks, Local Debt Risks and Shadow Banking Risks
	During the subprime crisis, shadow banking shifted risk short-term but did not disappear
	"Asset-liability maturity mismatch of some financial institutions represented by the "shadow banking system", increasing the vulnerability of the financial system
	Shadow banking, characterized by low transparency and complex structure, is prone to become a major hidden danger of systemic financial risk
Systemically important institutions (SIFIs)	The presence of "too big to fail" and "too connected to fail" large financial institutions in the financial system has a significant impact on systemic risk
Asset price volatility	An increase in the share of non-bank financial institutions can effectively take over assets sold at reduced prices after shocks to the banking system
	Asset price volatility as a source of financial fragility
Financial business structure	The business structure of financial institutions has a significant impact on systemic risk
	The use of non-interest income indicators to examine the impact of a financial institution's business income structure on its risks
Size of financial institutions	There is a significant positive association between the size of financial institutions and systemic risk
Scale of financial inclusion	Digital finance can curb systemic financial risks on a regional scale, but there is regional heterogeneity
	Digital financial inclusion can improve risk resilience within the financial system and can also facilitate the transfer of capital

Note: For reasons of space, this table shows only part of the data.

TABLE 2 Resilience indicator system for financial infrastructure.

Target layer	Standardized layer	Measurement indicators	Measurement equation	Resilience attribute	Indicator attributes	Weights
Resilience of financial infrastructure	Resilience of risk and control	Shadow banking risk	Shadow banking size/local and foreign currency loan balances	Impact	Negative	0.094
		Non-performing loan ratio	—	Impact	Negative	0.028
		Real estate liability risk	Real estate corporate debt Ratio	Impact	Negative	0.059
		Government debt risk	Local government debt/gdp	Impact	Negative	0.022
		Information asymmetry	Internet penetration rate	Resistance	Positive	0.112
	Resilience of recovery and stability	Level of financial inclusion	Digital inclusive finance index	Adjustment	Positive	0.29
		Financial regulatory intensity	Expenditures on financial regulation/general budget expenditures	Adjustment	Negative	0.05
		Insurance depth	Insurance income/gdp	Adjustment	Positive	0.087
	Resilience of transformation and development	Fintech level	Financial technology index	Transformation	Positive	0.166
		Industrial structure	Tertiary sector of industry/gdp	Transformation	Positive	0.092

Note: The standardized layer is derived from economic resilience theory; the measurement indicator sources, resilience attributes, and indicator attributes are derived from rootedness theory.

indicators will be assigned to one of the resilience attributes of impact, resistance, adjustment, and transformation, and the measurement indicators will be matched and corresponded with the guideline layer, which will ultimately constitute the financial infrastructure resilience indicator system as shown in Table 2.

Shadow banking is highly hidden, if directly measured and statistics are difficult, financial institutions do not have complete data [31–33]. Therefore, this paper is based on the borrower's point of view, drawing on the unobserved credit method to calculate the size of shadow banking in 31 provinces, which is expressed by the formula:

$$Size\,of\,shadow\,banking = \frac{L}{GDP} \times NOE$$

where L is the local and foreign currency loan balance of financial institutions in each province, and NOE denotes unobserved income in each province.

China's practice of financial inclusion and innovative digital finance show a strong correlation, and the new digital financial model has become an important source of impetus and growth for financial inclusion, the level of financial inclusion in this paper adopts the Peking University Digital Inclusive Finance Index. The current use of search engine data for modeling has become an effective and widely used research method [34–36], which provides insights into the use of Internet search to construct regional fintech development indicators. This paper utilizes Baidu search fintech related buzzword frequency to construct fintech index to reflect the development level of fintech.

4 Spatial-temporal dynamic evolution of provincial financial infrastructure resilience

4.1 Level of change in the resilience of provincial financial infrastructure

The changes in the resilience level of China's provincial financial infrastructure are shown in Figure 1. The average value of provincial financial infrastructure resilience increased from 0.159 in 2005 to 0.704 in 2020, showing an overall upward trend. Before 2008, the resilience level has been maintained at a low level. With the subprime crisis, China began to pay attention to the construction of financial infrastructure, and the level of resilience was better improved, however, it began to fall back to a certain extent from 2015.

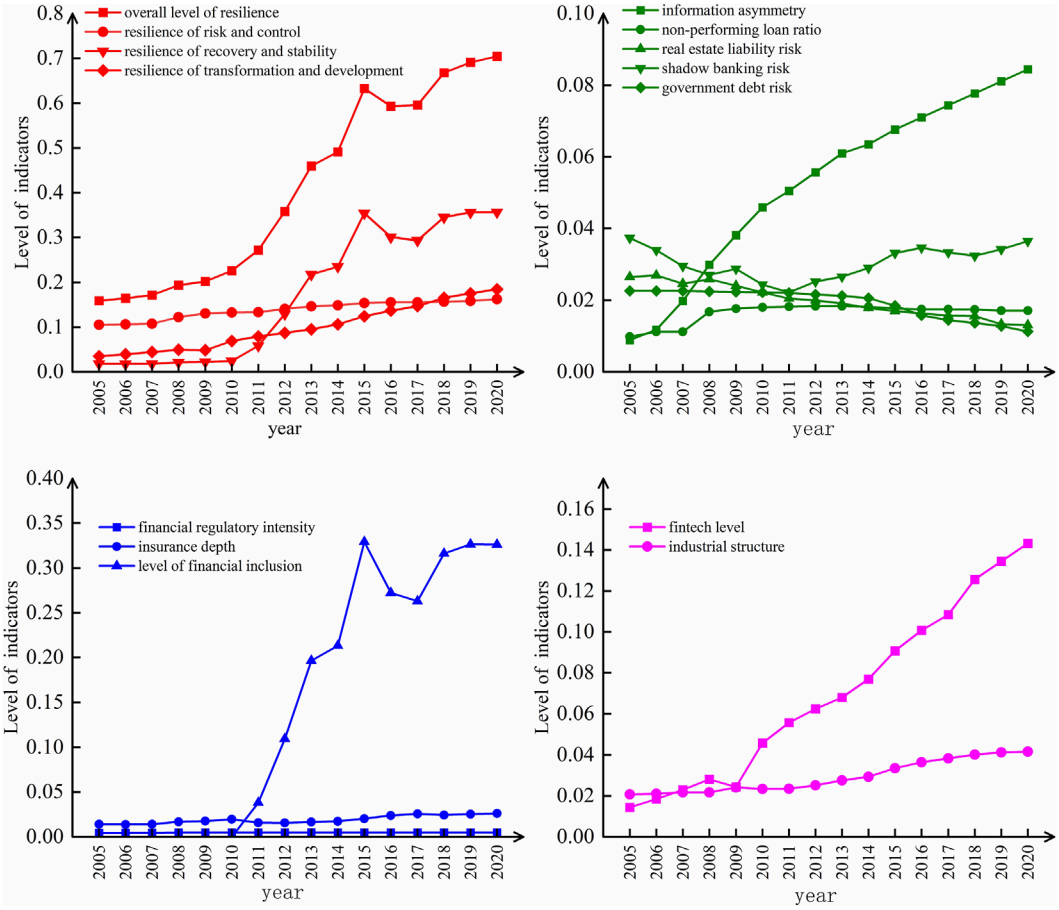


FIGURE 1
Resilience of China's provincial financial infrastructure and changes in indicators from 2005 to 2020.

Resilience of risk and control refers to the ability of financial infrastructure to maintain its normal operation in the face of adverse factors interference, as can be seen in Figure 1. Resilience of risk and control is in a state of stagnant growth, with the resilience value rising from 0.105 in 2005 to 0.162 in 2020. In terms of specific sub-indicators, the level of information asymmetry has been mitigated, which has contributed to the enhancement of the risk control resilience, but the shadow banking risk, real estate liability risk, and government debt risk have reduced resilience of risk and control to a certain extent.

Resilience of recovery and stability refers to the ability of the financial infrastructure to better cope with the external environment by adjusting its structure in the face of external disturbances, as can be seen in Figure 1, the resilience of recovery and stability has shown a clear upward trend, with the resilience value rising from 0.018 in 2005 to 0.35 in 2020. The insurance depth and the development of financial inclusion have better contributed to the resilience of recovery and stability, and financial inclusion in particular has played an important role in mitigating external shocks and supporting economic restructuring.

Resilience of transformation and development refers to the ability of financial infrastructure to generate innovative development after continuous adaptation and self-learning, Figure 1 shows that the resilience of transformation and development has shown a slight upward trend, with the resilience value rising from 0.035 in 2005 to 0.18 in 2020. Specifically, the increase in resilience of transformation and development is mainly due to the steady rise in fintech level, compared with the relatively limited impact of the industrial structure on resilience of transformation and development, and fintech level is the main driver of the increase in resilience of transformation and development.

4.2 Spatial-temporal trend in the resilience of provincial financial infrastructures

4.2.1 Trend in spatial variation of financial infrastructure resilience

Based on the global trend analysis method, as shown in Figure 2, this paper plots the spatial trend of financial infrastructure resilience in China's provincial areas over the entire sample period, as well as the 11th five-year plan, 12th five-year plan and 13th five-year plan periods, by setting the due east and due north directions as the X-axis and Y-axis, and the level of financial infrastructure resilience and the rate of change as the Z-axis.

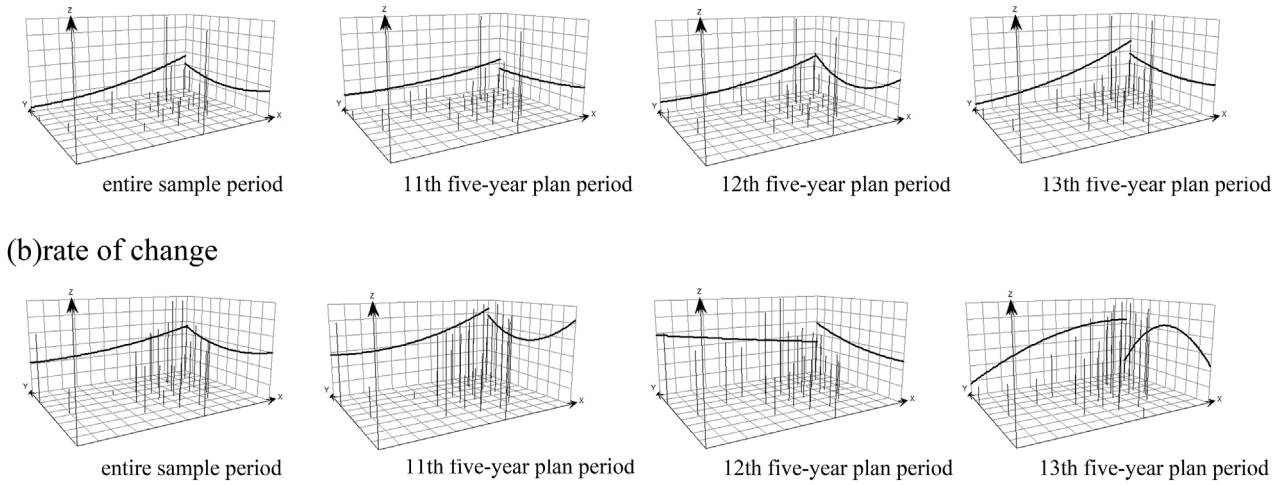


FIGURE 2
Trends in the spatial distribution of the level and rate of change of financial infrastructure resilience in provincial areas. **(A)** Resilience level. **(B)** Rate of change.

Over the entire sample period, as well as the 11th five-year plan, 12th five-year plan and 13th five-year plan periods, the spatial trend of financial infrastructure resilience show similar trends: The resilience level continues to rise from west to east, and shows a U-shaped trend from south to north. In terms of the spatial trend in the rate of change of financial infrastructure resilience, the entire sample period shows a trend of "low west and high east, north-south concave," but there are differentiation phenomena in different periods. Specifically, from an east-west perspective, the rate of change during the 12th five-year plan period showed a decreasing trend. From a north-south perspective, the 11th five-year plan period showed a U-shaped trend of change, the 12th five-year plan period showed a slight upward trend, and the 13th five-year plan period showed an inverted U-shaped trend of change.

Based on the above analysis, it can be concluded that there are significant spatial variations in the resilience level and rate of change of China's financial infrastructure, generally showing a "low west and high east, north-south concave" distribution trend, in which the spatial differentiation characteristics rate of rate of change is more obvious.

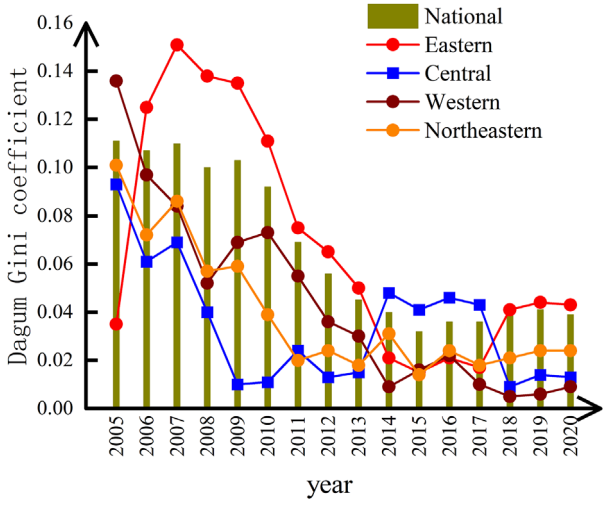


FIGURE 3
Dagum's gini coefficient for provincial financial infrastructure resilience from 2005 to 2020.

4.2.2 Spatial differentiation degree and sources of financial infrastructure resilience

The global trend analysis method reveals the absolute differences in the level of financial infrastructure resilience, but it is unable to describe the spatial and temporal evolution of the overall differences in the level of financial infrastructure resilience and the relative differences in China. Dagum Gini coefficient is able to measure the difference in resilience between different subsystems in the financial infrastructure, reflecting the vulnerability and imbalance of the financial system in the face of systemic risk. Based on this, this paper uses the Dagum Gini coefficient to analyze the numerical level of the overall differences in the resilience level of China's financial infrastructure and its sources, and identifies the changes in the relative differences among the four major economic zones, as shown in Figure 3.

The results show that, the Dagum Gini coefficient shows Eastern > Western > Northeastern > Central. The Dagum Gini coefficient in the western, northeastern and central regions generally shows a stepped downward trend, and the imbalance of financial infrastructure resilience within the region is narrowing; the Dagum Gini coefficient in the eastern region is significantly higher than other regions, and shows a trend of rising and then declining, which indicates that the resilience of the eastern provinces has a large gap between the level of the provinces, reflecting the obvious "wooden barrel effect". Therefore, the driving role of financial centers

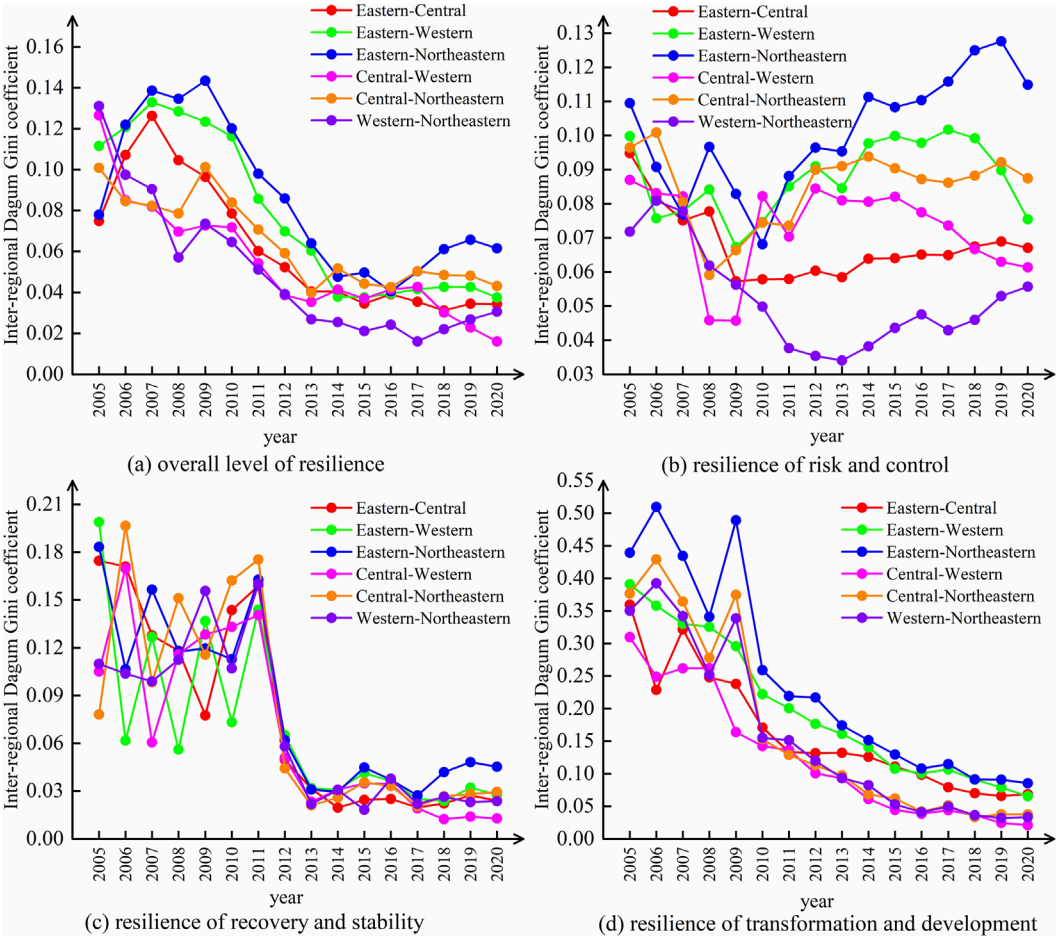


FIGURE 4
Interregional dagum gini coefficient for provincial financial infrastructure resilience from 2005 to 2020. **(A)** Overall level of resilience. **(B)** Resilience of risk and control. **(C)** Resilience of recovery and stability. **(D)** Resilience of transformation and development.

such as Shanghai and Shenzhen should be brought into play to promote in-depth synergies between their financial infrastructures and those of their neighboring provinces, to form a demonstration effect and a radiation effect, and to promote the sharing of resources and the exchange of technology. Secondly, policy guidance and financial support should be increased to promote the rational distribution of financial resources within the eastern region and avoid excessive concentration in some specific cities. In addition, financial institutions can be encouraged to play a more active role in cross-provincial business expansion through the establishment of regional financial cooperation platforms.

By applying the inter-regional Dagum Gini coefficient, this paper deeply explores the spatial differences in the resilience of provincial financial infrastructures among the four major economic zones of China (eastern, northeastern, central, and western) in terms of the overall level and the resilience of risk and control, the resilience of recovery and stability, and the resilience of transformation and development, as shown in Figure 4.

In terms of the overall level of resilience, the inter-regional differences roughly show Eastern-Northeastern > Eastern-Western > Central-Northeastern > Eastern-Central > Central-Western > Western-Northeastern, which indicates that Eastern lacks coordination with Northeastern and Western in the development of financial infrastructure resilience, Central has moderate differences with other regions, while Western and Northeastern have lower levels of financial infrastructure resilience and show similar development characteristics. Inter-regional differences show an overall downward trend, with a slight increase after 2017.

In terms of the resilience of risk and control, the inter-regional gap showed a trend of contraction several times from 2005 to 2010, indicating a gradual convergence in the ability of regions to cope with financial risks. However, the gap has been widening since 2010, especially between the eastern and the northeast, and between the eastern and the western, where the gap in risk resilience has become significantly wider. In terms of the resilience of recovery and stability, taking 2011 as the cut-off point, the inter-regional gap can be divided into two stages: before 2011, the gap was larger, and there were significant differences in the ability of regions to recover and stabilize financial services; after 2011, the gap was reduced and showed a wave-like trend of change. In terms of the resilience of transformation and development, except for the jump in inter-regional disparities in 2009, the overall trend is consistently

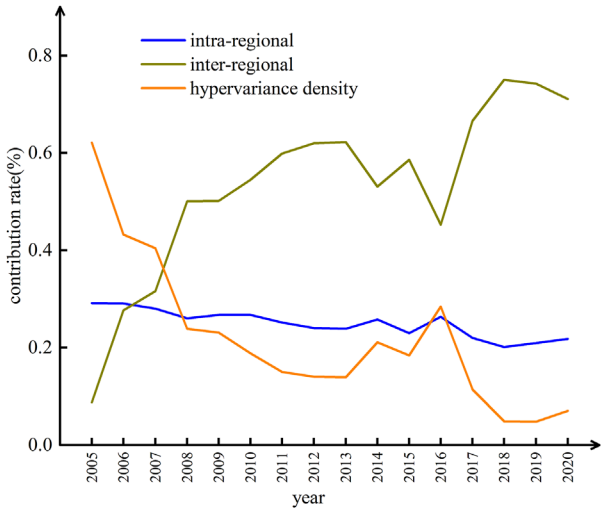


FIGURE 5
Regional gap contribution to provincial financial infrastructure resilience from 2005 to 2020.

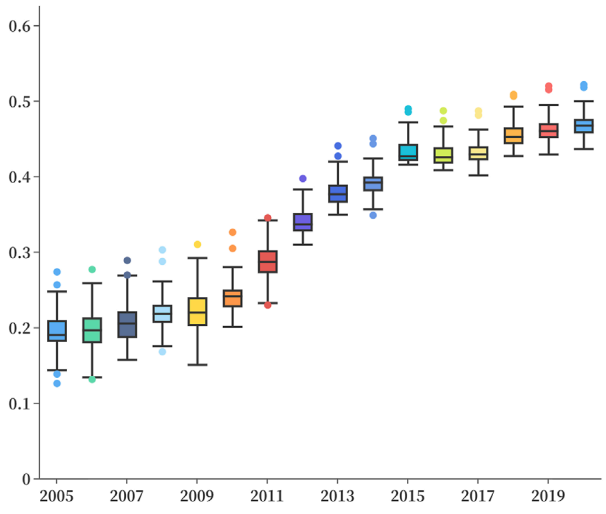


FIGURE 6
Boxplot of coupled harmonization of provincial financial infrastructure resilience.

downward, indicating that the differences in the ability of regions to cope with external changes through financial technology and innovation are gradually narrowing.

In terms of the contribution rate of regional disparities, as shown in Figure 5, the overall trend of inter-regional contribution rate is increasing, which indicates that in order to improve the overall level of provincial financial infrastructure resilience, more attention should be paid to inter-regional synergy and global coordinated development. The overall decreasing trend of inter-regional hypervariance density implies that the overlapping of inter-regional provincial samples also affects the spatial-temporal divergence of the resilience level, but the contribution rate of this effect is decreasing. In addition the intra-regional contribution rate consistently fluctuates between 20.14% and 29.13%, which is lower than the inter-regional contribution rate, but it should still be alerted to the divergence of the resilience level of provincial financial infrastructures within the four major economic regions.

4.3 Dynamic evolutionary characteristics of provincial financial infrastructure resilience

4.3.1 Coupling coordination analysis

The coupling coordination degree is used to portray the dynamic process of the resilience sub-dimension changing from disorder to order and forming a benign development. As shown in Figure 6, during the period from 2005 to 2020, the coupling coordination degree of the 31 provinces has a low degree of dispersion, fewer outliers, and a more consistent type of coupling coordination degree. The boxplot shows a wave-like increase over time, with the average value of coupling coordination degree of 0.193 in 2005 and 0.469 in 2020, and the level of coupling coordination is increasing. Influenced by economic, policy and other objective historical reasons, the financial infrastructure of Beijing and Shanghai enjoys a first-mover advantage, and the coupling coordination degree is always higher than that of other provinces. However, in terms of growth rate, the original low-coupling provinces have a more pronounced rate of increase. At present, the coupling and coordination level of provincial financial infrastructure resilience is still in the low-coupling stage, and there is still much room for improvement in the future.

Figure 7 shows the kernel density surface of the coupling coordination degree of financial infrastructure resilience for the period from 2005 to 2020. From the figure, it can be seen that the national kernel density curve oscillates to the right during the study period, indicating that the overall level of coupling coordination is improving; the height of the main peak of the kernel density curve is rising, the width is narrowing year by year, and the peak value of the main peak shows a trend of change from a "flat" to a "high" peak; the absolute gap and divergence trend of the coupling coordination level in most provinces are contracting; the right side of the kernel density curve shows a trend of contraction. The right tail of the kernel density curve shows a clear unimodal distribution, with the peak height continuously rising. The "bimodal club" characteristic is evident, indicating a polarization phenomenon in the resilience coupling coordination degree.

4.3.2 Overall evolution analysis

Spatial Markov chains are able to capture state changes and their interdependence at multiple levels of the financial infrastructure, especially when financial markets are subject to external shocks, and they can effectively reveal the dynamic evolution of different parts of the financial infrastructure and their risk transfer mechanisms. This paper constructs a Markov transfer probability matrix to analyze the overall evolution characteristics of China's financial infrastructure resilience. Table 3 is the provincial financial infrastructure resilience Markov transfer probability matrix for the period from 2005 to 2020, according to the calculation results show that: the probability value of the diagonal is greater than the probability value of the non-diagonal and more than 60%, which indicates that the

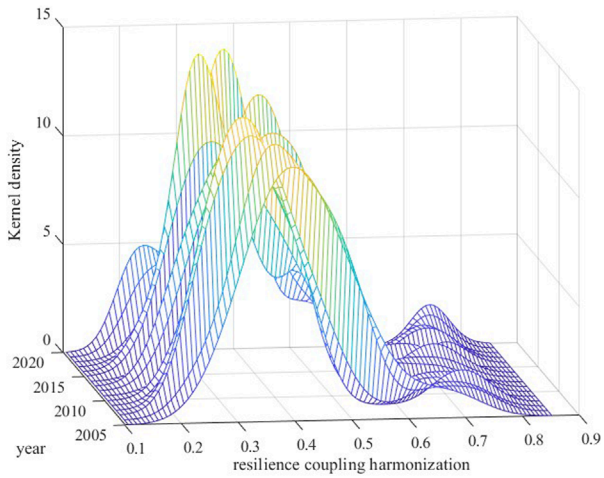


FIGURE 7
Kernel density surface map of coupled coordination degree of financial infrastructure resilience.

TABLE 3 Traditional markov transfer probability matrix for financial infrastructure resilience.

t/t+1	n	1	2	3	4
1	124	0.734	0.266	0.000	0.000
2	124	0.040	0.710	0.250	0.000
3	124	0.000	0.000	0.629	0.371
4	93	0.000	0.000	0.161	0.839

TABLE 4 Spatial markov transfer probability matrix for financial infrastructure resilience.

Types of spatial lag	t/t+1	n	1	2	3	4
1	1	107	0.822	0.178	0.000	0.000
	2	27	0.185	0.815	0.000	0.000
	3	0	0.000	0.000	0.000	0.000
	4	0	0.000	0.000	0.000	0.000
2	1	17	0.176	0.824	0.000	0.000
	2	94	0.000	0.702	0.298	0.000
	3	2	0.000	0.000	1.000	0.000
	4	0	0.000	0.000	0.000	0.000
3	1	0	0.000	0.000	0.000	0.000
	2	3	0.000	0.000	1.000	0.000
	3	103	0.000	0.000	0.631	0.369
	4	10	0.000	0.000	0.300	0.700
4	1	0	0.000	0.000	0.000	0.000
	2	0	0.000	0.000	0.000	0.000
	3	19	0.000	0.000	0.579	0.421
	4	83	0.000	0.000	0.145	0.855

transfer of the level of financial infrastructure resilience type has the stability of the transfer, and maintain the original state of the probability of greater. There is a "club convergence" phenomenon in the evolution of China's financial infrastructure resilience, and the probability that the low resilience and high resilience types will maintain the original state type in the next stage is the greatest, 73.4% and 83.9%, respectively. It is difficult to realize the "leapfrog" development of the transfer of resilience level types in adjacent years, and the probabilities of the low-resilience level, medium-low-resilience level, and medium-high-resilience level jumping up one level are 0.26, 0.25, and 0.37, respectively, which indicate that there is the possibility of upward development of all three levels. Overall, the results confirm the trend of upward development of financial infrastructure resilience levels in various provinces of China over the years, as well as the relative stability of development.

On the basis of the traditional Markov transfer probability matrix, the spatial lag condition is introduced to construct the spatial Markov transfer probability matrix, aiming at exploring the influence on the transfer probability of the resilience level type in different spatial lag backgrounds. According to the calculation results in Table 4, it is shown that spatial elements have influence in the evolution of financial infrastructure resilience type, when a province is adjacent to a neighbor with a low level of resilience, the probability of the province's resilience level type transferring upward decreases, and when adjacent to a neighbor with a high level of resilience, the probability of the province's resilience level type transferring upward will increase. For example, when in the condition of a neighborhood with a low resilience level, $P_{11|1}(0.822)>P_{11}(0.734)$, $P_{12|1}(0.178)<P_{12}(0.266)$, while in the condition of a neighborhood with a high resilience level, $P_{34|4}(0.421)>P_{34}(0.371)$, $P_{44|4}(0.855)>P_{44}(0.839)$. The spatial Markov transfer probability matrix provides an explanation for the "club convergence" phenomenon in the spatial dimension. Influenced by the spillover effect of neighborhood types, the transfer of financial infrastructure resilience types is prone to form the phenomenon of "club convergence" within a certain geospatial range.

4.4 Barrier factor analysis

After analyzing the coupling coordination and overall evolution of provincial financial infrastructure, further research to identify the barrier factors influencing the dynamic changes in resilience levels can not only provide a deeper understanding of the

TABLE 5 Main barrier factors to the resilience of provincial financial infrastructure from 2005 to 2020.

2005		2010		2015		2020	
Barrier factors	Barrier degree	Barrier factors	Barrier degree	Barrier factors	Barrier degree	Barrier factors	Barrier degree
Level of financial inclusion	14.062	Level of financial inclusion	15.309	Fintech level	6.362	Shadow banking risk	6.335
Fintech level	5.588	Fintech level	4.807	Shadow banking risk	5.199	Level of financial inclusion	5.616
Information asymmetry	3.802	Shadow banking risk	2.815	Industrial structure	4.942	Industrial structure	5.301
Industrial structure	2.64	Industrial structure	2.759	Level of financial inclusion	4.403	Real estate liability risk	5.05
Shadow banking risk	2.098	Information asymmetry	2.627	Information asymmetry	3.68	Information asymmetry	2.78
Real estate liability risk	1.229	Real estate liability risk	1.512	Real estate liability risk	3.657	Fintech level	2.255
Insurance depth	1.218	Insurance depth	1.107	Insurance depth	2.267	Insurance depth	2.236
Non-performing loan ratio	0.332	Non-performing loan ratio	0.033	Government debt risk	0.359	Government debt risk	1.201
Financial regulatory intensity	0.031	Government debt risk	0.02	Non-performing loan ratio	0.095	Non-performing loan ratio	0.18
Government debt risk	0	Financial regulatory intensity	0.011	Financial regulatory intensity	0.035	Financial regulatory intensity	0.047

characteristics of resilience level changes but also offer strong support for formulating financial system policies and behavioral adjustments, thereby effectively improving the resilience of financial infrastructure. This paper uses barrier degree model to measure the barrier degree of factors affecting financial infrastructure resilience from 2005 to 2020.

From the perspective of the barrier degree of individual indicators as shown in Table 5, the barrier degree of the proportion of industrial structure, shadow banking risk, real estate liability risk, and fintech level has risen significantly, while the barrier degree of the level of financial inclusion and information asymmetry has shown a downward trend, and the barrier degree of the intensity of financial regulatory intensity has not fluctuated significantly.

In terms of the barrier degree of the financial infrastructure resilience sub-dimension as shown in Figure 8, the barrier degree of resilience of risk and control is on an upward trend, indicating that the ability of the financial infrastructure to resist unexpected external shocks is weakening. The downward trend in the resilience of recovery and stability suggests that after years of development, the adaptive capacity of the financial infrastructure to cope with crises is increasing. The barrier degree of resilience of transformation and development has not fluctuated much, suggesting that the financial infrastructure's momentum in innovative development is relatively weak.

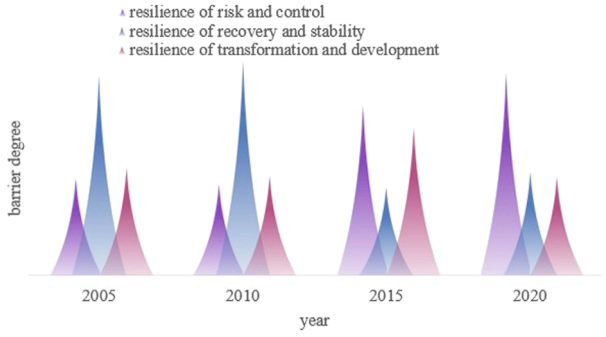


FIGURE 8
Provincial financial infrastructure resilience subsystem Barriers from 2005 to 2020.

5 Conclusion

This paper takes the 31 provinces of China from 2005 to 2020 as the research units, constructs an evaluation index system for provincial financial infrastructure, and measures the resilience levels. On this basis, the Dagum Gini coefficient, spatial Markov chain, and barrier degree are used to analyze the variation level, spatial-temporal differentiation, and dynamic evolution

characteristics of provincial financial infrastructure resilience. The main conclusions are as follows.

(1) the resilience level of financial infrastructure across China's provinces shows an overall upward trend, but a decline in resilience is observed after 2015. From the perspective of resilience sub-dimensions, resilience of risk and control is in a state of slow growth, resilience of recovery and stability shows a significant upward trend, while resilience of transformation and development exhibits little fluctuation and no clear upward trend.
(2) In terms of spatial-temporal differentiation characteristics, the level of financial infrastructure resilience and the rate of change roughly shows "low west and high east, north-south concave" distribution trend, the rate of change of the spatial differentiation characteristics are more obvious. Dagum Gini coefficient shows that the East > West > Northeast > Central, the eastern region of the financial infrastructure resilience level gap is larger, the "wooden barrel effect" is obvious. The inter-regional contribution rate is much higher than the intra-regional contribution rate and hypervariable density, and is on the rise. The hypervariable density contribution rate shows a decreasing trend, which means that the impact of interregional provincial samples overlapping and thus causing differences in resilience levels is decreasing. The intra-regional contribution rate maintains a small fluctuation between 20.14% and 29.13%, which should alert to the divergence of the resilience level of the financial infrastructure of the provincial regions within the region.
(3) From the perspective of dynamic evolution characteristics, the coupling coordination degree of resilience sub-dimensions is rising in a wave-like manner, but is still in the medium-low coupling stage. The transfer of the type of financial infrastructure resilience level in China is stable, influenced by the spillover effect of the type of neighborhood, and prone to the formation of the "club convergence" phenomenon within a certain geospatial range.
(4) From the perspective of barrier factors, there is a significant increase in the barrier degree of industrial structure, shadow banking risk, real estate liability risk, and fintech level. Resilience of risk and control, resilience of recovery and stability and resilience of transformation and development barrier degree show a rising, falling and stable trend of change, respectively.

Although this study thoroughly explores the evolutionary characteristics of provincial financial infrastructure resilience and its influencing factors, there are still some limitations. First, limited by data availability, this paper mainly considers the frequency and accessibility of data when constructing the indicator system, which may affect the precision of the resilience assessment. In addition, the study does not fully consider the short-term impact of sudden external shocks on the resilience of financial infrastructures, and the role of these factors can be further explored in the future. Although fintech is considered an important influence on resilience, this study fails to quantify its specific impact, and the differentiated role of fintech among different provinces could be analyzed in depth in the future. The above limitations provide room for improvement and direction for future research.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: The data are not publicly available. Requests to access these datasets should be directed to zhongwei_86@163.com.

Author contributions

GL: Writing–review and editing. ZX: Writing–original draft. CK: Writing–review and editing. SL: Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (Grant Nos 72374063 and 72304089).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. León C, Berndsen R, Renneboog L. Financial stability and interacting networks of financial institutions and market infrastructures. European Banking Center Discussion Paper Series, (2014-011) (2014).

2. Li F, Perez-Saiz H. Measuring systemic risk across financial market infrastructures. *J Financial Stab* (2018) 34:1–11. doi:10.1016/j.jfs.2017.08.003

3. Martin R. Regional economic resilience, hysteresis and recessionary shocks. *J Econ Geogr* (2012) 12(1):1–32. doi:10.1093/jeg/lbr019

4. Berndsen R, Heijmans R. Risk indicators for financial market infrastructure: from high frequency transaction data to a traffic light signal (2017).

5. Berndsen RJ, León C, Renneboog L. Financial stability in networks of financial institutions and market infrastructures. *J Financial Stab* (2018) 35:120–35. doi:10.1016/j.jfs.2016.12.007

6. Wendt F. *Central counterparties: addressing their too important to fail nature*. International Monetary Fund (2015).

7. Lee R. *Running the world's markets: the governance of financial infrastructure*. Princeton University Press (2010).

8. Peter Gray H. The quality of financial infrastructure and financial resilience: contrasting Taiwan and Thailand. *Rev Pac Basin Financial Markets Policies* (2002) 5(02):181–94. doi:10.1142/s0219091502000778

9. Dupont B. The cyber-resilience of financial institutions: significance and applicability. *J cybersecurity* (2019) 5(1):tyz013. doi:10.1093/cybsec/tyz013

10. Costabile L, Neal L (2018). Financial innovation and resilience. A comparative perspective on the public banks of naples (1462-1808).

11. Kass-Hanna J, Lyons AC, Liu F. Building financial resilience through financial and digital literacy in South Asia and Sub-Saharan Africa. *Emerging Markets Rev* (2022) 51:100846. doi:10.1016/j.ememar.2021.100846

12. Salignac F, Marjolin A, Reeve R, Muir K. Conceptualizing and measuring financial resilience: a multidimensional framework. *Soc Indicators Res* (2019) 145:17–38. doi:10.1007/s11205-019-02100-4

13. Zhang L, Ma X, Ock YS, Qing L. Research on regional differences and influencing factors of Chinese industrial green technology innovation efficiency based on dagum gini coefficient decomposition. *Land* (2022) 11(1):122. doi:10.3390/land11010122

14. Ma T, Liu Y, Yang M. Spatial-temporal heterogeneity for commercial building carbon emissions in China: based the Dagum Gini Coefficient. *Sustainability* (2022) 14(9):5243. doi:10.3390/su14095243

15. Hong H, Liao H, Li T, Yang J, Xie D. Analysis of spatio-temporal patterns of rural space function based on entropy value method and Dagum Gini coefficient. *Trans Chin Soc Agric Eng* (2016) 32(10):240–8. doi:10.11975/j.issn.1002-6819.2016.10.034

16. Zhou R, Jin J, Cui Y, Ning S, Zhou L, Zhang L, et al. Spatial equilibrium evaluation of regional water resources carrying capacity based on dynamic weight method and Dagum Gini coefficient. *Front Earth Sci* (2022) 9:790349. doi:10.3389/feart.2021.790349

17. Kakamu K. Simulation studies comparing Dagum and Singh–Maddala income distributions. *Comput Econ* (2016) 48:593–605. doi:10.1007/s10614-015-9538-z

18. Carle SF, Fogg GE. Modeling spatial variability with one and multidimensional continuous-lag Markov chains. *Math Geology* (1997) 29:891–918. doi:10.1023/a:1022303706942

19. Dehghan Shabani Z, Shahnazi R. Spatial distribution dynamics and prediction of COVID-19 in Asian countries: spatial Markov chain approach. *Reg Sci Pol and Pract* (2020) 12(6):1005–26. doi:10.1111/rsp3.12372

20. Shepero M, Munkhammar J. Spatial Markov chain model for electric vehicle charging in cities using geographical information system (GIS) data. *Appl Energ* (2018) 231:1089–99. doi:10.1016/j.apenergy.2018.09.175

21. Eaker MR, Crane DB, Froot KA, Mason SP, Perold AF, Merton RC, et al. The global financial system: a functional perspective. *J Finance* (1997) 52:915. doi:10.2307/2329507

22. King RG, Levine R. Finance and growth: schumpeter might be right. *Q J Econ* (1993) 108(3):717–37. doi:10.2307/2118406

23. Arestis P, Demetriades P. Financial development and economic growth: assessing the evidence. *Econ J* (1997) 107(442):783–99. doi:10.1111/j.1468-0297.1997.tb00043.x

24. Martin R, Sunley P. On the notion of regional economic resilience: conceptualization and explanation. *J Econ Geogr* (2015) 15(1):1–42. doi:10.1093/jeg/lbu015

25. Simmie J, Martin R. The economic resilience of regions: towards an evolutionary approach. *Cambridge J regions, economy Soc* (2010) 3(1):27–43. doi:10.1093/cjres/rsp029

26. Hill E, Wial H, Wolman H (2008). Exploring regional economic resilience.

27. Pistor K. A legal theory of finance. *J Comp Econ* (2013) 41(2):315–30. doi:10.1016/j.jce.2013.03.003

28. Wolfswinkel JF, Furtmueller E, Wilderom CP. Using grounded theory as a method for rigorously reviewing literature. *Eur J Inf Syst* (2013) 22(1):45–55. doi:10.1057/ejis.2011.51

29. Eisenhardt KM. Building theories from case study research. *Acad Manag Rev* (1989) 14(4):532–50. doi:10.2307/258557

30. Fernandez WD, Lehmann H. Case studies and grounded theory method in information systems research: issues and use. *J Inf Technology Case Appl Res* (2011) 13(1):4–15. doi:10.1080/15228053.2011.10856199

31. Barth JR, Li T, Shi W, Xu P. China's shadow banking sector: beneficial or harmful to economic growth? *J Financial Econ Pol* (2015) 7(4):421–45. doi:10.1108/jfep-07-2015-0043

32. Chen K, Ren J, Zha T. The nexus of monetary policy and shadow banking in China. *Am Econ Rev* (2018) 108(12):3891–936. doi:10.1257/aer.20170133

33. Allen F, Qian Y, Tu G, Yu F. Entrusted loans: a close look at China's shadow banking system. *J Financial Econ* (2019) 133(1):18–41. doi:10.1016/j.jfineco.2019.01.006

34. Xue G, Liu S, Ren L, Gong D. Forecasting hourly attraction tourist volume with search engine and social media data for decision support. *Inf Process and Management* (2023) 60(4):103399. doi:10.1016/j.ipm.2023.103399

35. Tian F, Yang Y, Mao Z, Tang W. Forecasting daily attraction demand using big data from search engines and social media. *Int J Contemp Hospitality Management* (2021) 33(6):1950–76. doi:10.1108/ijchm-06-2020-0631

36. Segev E. Googling the world: global and regional information flows in Google Trends. *Int J Commun* (2018) 12:19.

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 03 March 2025
DOI 10.3389/fphy.2025.1559799

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Xiaoqing Hao,
Nanjing University of Aeronautics and Astronautics, China
Zhihua Chen,
Beijing Normal University, China

*CORRESPONDENCE
Kun Wang,
wkun1111@126.com

RECEIVED 13 January 2025
ACCEPTED 06 February 2025
PUBLISHED 03 March 2025

CITATION
Chen Q, Wang K, Zhang Y, Guan Q, Xing J, Long T, Zheng G, Li Q, Li Z, Ren X, Shang C and Duan Y (2025) Simulation analysis on the evolution driven model of global copper ore trade network.
Front. Phys. 13:1559799.
doi: 10.3389/fphy.2025.1559799

COPYRIGHT
© 2025 Chen, Wang, Zhang, Guan, Xing, Long, Zheng, Li, Li, Ren, Shang and Duan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Simulation analysis on the evolution driven model of global copper ore trade network

Qishen Chen[1], Kun Wang[1]*, Yanfei Zhang[1], Qing Guan[2], Jiayun Xing[1], Tao Long[1], Guodong Zheng[1], Qiang Li[1], Zhenqing Li[1], Xin Ren[1], Chenghong Shang[1] and Yueran Duan[2]

[1]Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing, China, [2]School of Information Engineering, China University of Geosciences (Beijing), Beijing, China

Introduction: In the context of energy transition, the competition for copper resources among countries has intensified, and the global copper trade has become a vitally important trade chain. The global copper ore trade network is influenced by various factors, including resource distribution, supply, demand, prices, transportation costs, etc.

Methods: To understand the evolution process of copper trade network and to predict the trend of supply chain structure evolution in future, in this paper, we construct a spatial weighted complex network evolution model based on complex network theory and gravity model using the import and export data and distance data of countries from 1990 to 2022.

Results and discussion: Simulation results show that the possibility of establishing copper ore trade between countries follows the spatial weighted complex network evolution model. It is proportional to the expected trade flow between countries and inversely proportional to the distance. The model will support the simulation analysis of the supply chain network structure evolution and help to carry out in-depth research on the forecast of future trade relations between important countries.

KEYWORDS

spatial complex network, gravity model, copper ore supply chain, network evolution pattern, dynamics simulation

1 Introduction

In recent years, copper has been widely used in new energy-related industries, such as transmission lines, transformers, cable shielding belts and high-end radiators, which are important support for the construction of renewable energy generation installations (such as wind turbines) and solar photovoltaic panels [1–3]. With the rapid development of clean energy and energy-saving technologies, and the global copper demand will continue to grow in a long term [4–6]. Copper resources are abundant globally, but the distribution is unbalanced among countries, due to the different geological conditions [7]. Major copper exporters include Chile, Peru and the Democratic Republic of Congo, and major importers are concentrated in Asia and Europe, such as China, Japan, Germany, Spain and so on. With the increase of copper demand, the competition for copper resources among countries has intensified, and the global copper trade has become a vitally important trade chain. Therefore,

studying the supply chain evolution pattern of copper with different characteristics in international trade helps to study the formation and evolution mechanism of copper trade pattern, and predict the trend of supply chain structure evolution based on future trade demand.

In recent years, the study of mineral resources supply chain and the associated risks have gradually become a current research Frontier and hot spot in the field of resources, which has received attention from global scholars [8–12]. The research object gradually expands from energy minerals such as oil, natural gas to emerging minerals such as lithium, nickel and rare earth involved in new energy and other emerging industries [13–15]. The research mainly focuses on the trade flows among countries, trade structure characteristics of different mineral products in the whole industry chain and risk resistance of supply chain network [16]. Based on the industry chain perspective, many scholars have made a series of research advances by constructing models to research structure, resilience and risk of supply chain by complex network, SD simulation, material flow analysis (MFA/SFA), gravity model, Markov Chain and other simulation models [17–23]. Complex network can systematically describe the structural characteristics of mineral resource supply chains. Many Scholars have established network models based on trade flows, taking countries as nodes [24, 25]. Most of them focus on the description of topological characteristics of trade networks, the analysis of important nodes and the propagation of risks in supply chain networks [26, 27]. Fewer studies have been conducted on the driving models and mechanisms behind the evolution of supply chain network structures. Some research considers the influence of a single factor or multiple factors such as supply, demand, price, economics and others on the global copper trade using the gravity model, but little evidence has been provided to reveal the evolutionary mechanism of the copper trade network structure, especially the role of geographical factors in the evolution processed of network structures.

Therefore, this paper proposes a spatial weighted complex network evolution model that integrates the complex network theory and gravity model to analyze the evolutionary mechanism of network structure in copper trade. Utilizing this model facilitates the simulation and examination of the evolutionary dynamics within the spatially weighted network of the global copper trade. It explains the formation of the current trade network structure and reveals the role that geographical factors play in shaping trade patterns. The import and export data and distance data of countries from 1990 to 2022 are used to simulate the evolution process of the global copper network. The parameters of the model are discussed by comparing the simulated network structure with the real network structure. Based on the model, the trend of supply chain structure and the potential trading countries can be forecasted. The approach is helpful for trading countries to master their own characteristics and development trends of trade structure, and prepare for trade risks in a timely manner to reduce trade risks. The remainder of this paper is organized as follows. Section 2 describes the construction of the spatial weighted complex network evolution model and data sets used in our work. Section 3 presents the results about evolutionary process of the global copper ore trade network. We summarize our conclusions in Section 4.

2 Methods and data

2.1 Complex network model of international trade in copper ore

According to the complex network theory, this paper constructs a weighted directed network model for the international copper ore trade, as shown in Equation 1 [28]:

$$G = (N, E, W) \tag{1}$$

where N is the node, which is the set of all the countries involved in copper ore trade, E is the edge between two nodes. It represents the trade relationship between two country nodes, and the direction of the edge represents the trade flow between two nodes. If there is a trade relationship between two country nodes, there is a continuous edge between the two nodes; if there is no trade relationship between the two country nodes, there is no continuous edge between the two nodes. W is the weight of the continuous edge, which is the trade volume between the two country nodes.

The digraph G is represented by a matrix φ_{ij} with N rows and N columns. If there is a directed edge between node i and node j, and its weight is τ, as shown in Equation 2.

$$\varphi_{ij} = \varphi[i,j] = \tau \tag{2}$$

The average path length signifies the mean trade distance among countries, with shorter lengths indicating closer and more intimate trade relationships between them [29]. The definition of average short path length(L) is given in Equation 3.

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \tag{3}$$

where N is the total number of nodes, d_{ij} is the shortest path between the nodes i and j.

The average clustering coefficient serves as a metric to quantify the proximity of trading partners in terms of their trade structure [22]. A higher average clustering coefficient implies a tighter relationship, thereby enhancing the likelihood of establishing long-term and stable trade relations among the partners. The definition of the average clustering coefficient of a network is given in Equation 4.

$$C = \frac{1}{N} \sum_{i=1}^{N} \frac{n_i}{g_i(g_i - 1)} \tag{4}$$

where C represents the average clustering coefficient, n_i is the number of edges between all adjacent nodes of node i; g_i is the number of all neighboring nodes of node i.

The betweenness centrality of a node is often used to reflect its contribution to network connectivity [30]. It is calculated as the ratio of the number of shortest paths passing through the node to the total number of all shortest paths in the network. The larger the BC_i, the stronger the intermediary regulation ability, and the stronger the control over the trade network. The definition of the betweenness centrality (B_i) of a node is given in Equation 5.

$$B_i = \frac{2}{(N-1)(N-2)} \sum_{p=1}^{N} \sum_{q=1}^{N} \frac{R_{pq(i)}}{R_{pq}} \tag{5}$$

where R_{pq} is the shortest path between p and q, $R_{pq(i)}$ indicating that there is the shortest path from p to q through node i ($i \neq p \neq q$; $p = 1, 2, 3, \ldots, N; q = 1, 2, 3, \ldots, N$).

The closeness centrality measures the proximity of a trading country to other countries within a complex network, typically quantified as the reciprocal of the summed shortest paths between that country and all others. The higher the value of closeness centrality C signifies the trading country is more significant and difficult to be controlled by other countries [31]. Closeness centrality (K_i) is defined by Equation 6.

$$K_i = \frac{1}{\sum_{j=1}^{N} d_{ij}} \tag{6}$$

The eigenvector centrality uses the importance and influence of neighboring countries to describe the role of the country's trade in the complex network [30]. These countries display the attribute of "strengthening through collaboration with prominent trade nations," implying that countries with higher eigenvector centrality are typically interconnected with major trading hubs. This underscores their capacity to exert indirect influence on the trade structure. Eigenvector centrality (μ_i) is defined by Equations 7, 8.

$$HX = \mu X \tag{7}$$

$$\mu_i x_i = h_{1i}x_1 + h_{2i}x_2 + \cdots + h_{Ni}x_i \tag{8}$$

where H is an $N \times N$ adjacency matrix composed of h_{ij}. h_{ij} indicates the importance measure value of node i to j. $x = x_1, x_2, \cdots x_N$ indicates the degree centrality of each node. μ_i denotes the eigenvector centrality of node i.

2.2 Spatial weighted complex network evolution based on gravity model

Improved gravity models have been widely used to estimate and study inter-city population migration, traffic flow or trade flow [32, 33]. In actual networks such as air and road networks, the influence of geographical factors on the network leads to the "convergence" of network edges [34]. The evolution of the global copper supply chain structure is influenced by a combination of factors such as resource endowments, supply and demand patterns, market prices, geopolitical relationships, regional trade agreements, and transportation costs. The transportation cost can be reflected by the geographical distance between two locations to a certain extent. This paper constructs a spatial weighted complex network model of international trade in copper ore, from the perspective of the dynamics of the evolution of the network structure. The supply-demand relationship and transportation costs of each country node in constructing trade relations have been taken into account. Furthermore, this study has delved into the characteristics of the evolution of the international trade structure for copper ore, along with an examination of the pivotal role played by geographical factors in this evolutionary process.

In this paper, we introduce an improved gravity model as a measure of expected trade flows between nodes, defined by Equation 9.

$$F_{ij} = S\frac{m_i^{\alpha} m_j^{\alpha}}{r_{ij}^{\varepsilon}} \tag{9}$$

where F_{ij} is the expected trade flow between nodes, S is a constant, m_i and m_j represent the degree of adaptation of node i and node j. It is defined as the total export volume of exporting countries and the total import volume of importing countries, representing the trade demand of the country. In general, nodes with larger adaptability often have larger degree values. In this study, the "adaptability" of a node represents the trade demand of a country. A larger value of m_i or m_j indicates a greater trade demand, and thus the node is more likely to establish trade connections, resulting in a larger degree value. r_{ij} is the distance between two countries. Since there are often multiple transportation paths between two countries, it is difficult to obtain actual distance data. Normally, the capital is the economic and trade center of the country, and the distance r_{ij} mainly represents the transportation costs between two countries. So this paper simplifies the transportation distance as a straight-line distance between two national capitals, which is the Euclidean distance between two nodes. Although the actual distance among some countries may be biased, it can roughly reflect the transportation costs in the global iron ore trade. α and ε are two parameters, and their values depend on the network's ability for nodes to adaptability and dependence on geographical factors. α as the adaptation factor, represents the influence of trade demand on network formation; and ε as the geographical factor, captures the role of distance in trade resistance. It has been proved that in actual networks such as airline networks and urban road networks, $\alpha = 1$ and ε fluctuates between 0.2 and 2.7 [35–39].

Using the improved gravity model described above, it is possible to predict the trade volume between nodes, describing potential dynamics when the nodes are not yet connected to each other. Since F_{ij} describes only the trade volume generated between node i and node j, it does not include those generated by other node pairs that propagate through some non-direct path through the (i, j) edge.

Based on the above formula the expected trade flow between any two nodes can be calculated. According to the actual network development law, when the cost is limited, investment will be preferentially made in the construction of the areas with the most urgent needs and the largest trade flow demands. Only in this way can the dynamic needs of the network be met, and the expected returns of the network be maximized, as shown in the Equation 10.

$$\begin{aligned} Max W_F &= \sum_{i<j} F_{ij}\varphi_{ij} = \sum_{i<j} S\frac{m_i^{\alpha} m_j^{\alpha}}{r_{ij}^{\varepsilon}}\varphi_{ij} \\ &\sum_{i<j} \varphi_{ij} = \delta \end{aligned} \tag{10}$$

where φ_{ij} is the adjacency matrix of the network, δ is the set number of network edges, and W_F is the total expected trade flow of the network. In building the network, priority is given to connecting those node pairs with larger expected trade flows, starting from the pair of nodes with the largest F_{ij}, and connecting the corresponding node pairs in decreasing order of F_{ij} until the simulation network reaches a certain number of edges.

2.3 Data sources

Because the import and export data are collected in different countries, there are some differences in the import and export data reported by countries. For example, the trade volume of China's exports to Japan comes from China, while the trade volume of Japan's imports from China comes from Japan. Therefore, the base data used in this paper are from the copper ore import data of UN Comtrade website for the years 1990, 1995, 2000, 2005, 2010, 2015, 2020 and 2021. The copper ore data name is "Copper ores and concentrates," HS number 2603. Especially the 2021 data is used as the benchmark for the simulation of the supply chain network evolution process. The 2021 data contains a total of 117 countries or regions. Some statistical objects such as "other regions of the world, other regions of Asia" cannot be accurately located to the actual geographical location, so they are excluded in the process of this study. After screening, a total of 114 nodes with 580 connected edges were involved in the modeling and analysis of the network evolution process. For ease of display, the abbreviations of country names are used in the figure.

The data of the distance between countries were obtained from the CEPII database. The CEPII database is a comprehensive multinational database, which contains a large number of trade and globalization related data from various countries and regions around the world. The database is widely used in the field of trade and globalization research and has become a reliable source of theoretical and empirical research for many research institutions, governments and academia worldwide. The geographical distance data between two countries is an indicator provided by the cep database. To facilitate the comparison between models, the trade volume and distance data are normalized in this paper, and the raw data are transformed to the interval [1,10].

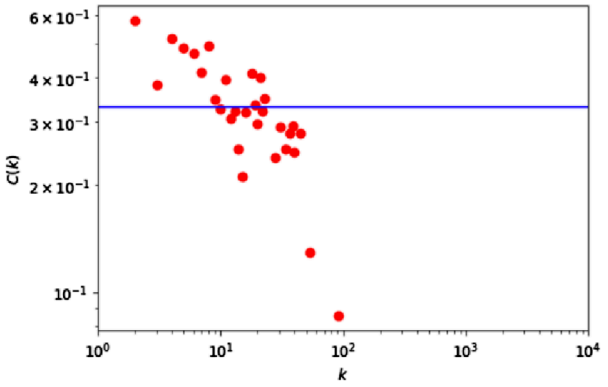


FIGURE 1
Distribution of model clustering coefficient-degree correlation.

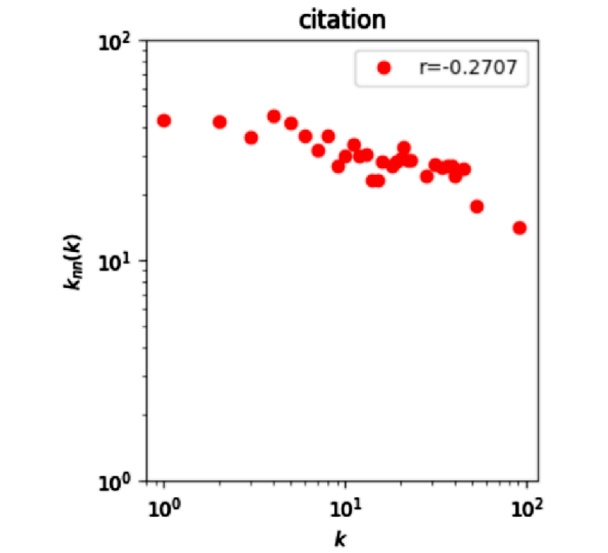


FIGURE 2
Network model degree - degree correlation distribution.

3 Modeling and results

3.1 Parametric analysis of global copper ore trade network model

In 2021, the global copper ore trade network contains 114 country nodes with 580 connected edges, including 99 nodes in exporting countries and 72 nodes in importing countries. Through the clustering coefficient C(k)-degree(k) correlation distribution characteristics of the network model (Figure 1), the dependence between the clustering coefficient and degree of the network model can be judged. The clustering coefficient C(k) of the network model decreases with the increase of node degree, so the network is a small-world network. With a small average path length of 2.54 and a small network density of 0.045, it indicates that the countries in this network are loosely connected. An average clustering coefficient value of 0.324 suggests that nodes within the network have a propensity to form densely interconnected clusters, indicating a higher degree of local clustering compared to that observed in random networks [27].

The degree-degree correlation of the network model characterizes the relationship between nodes with large degree and nodes with small degree in the network. If nodes with large degree tend to connect with nodes with large degree, the network exhibits a positive degree-degree correlation feature. Conversely, nodes with large degree tend to connect with nodes with small degree, the network exhibits a negative degree-degree correlation feature. According to the degree-degree correlation distribution of the network model (Figure 2), it can be seen that the nearest neighbor average degree value (knn(k)) of the nodes with degree k is a decreasing function that rises with k. This indicates that the network is a heterogeneous network with negative correlation characteristics, and nodes with large degree tend to connect with nodes with small degree.

By constructing the global copper ore trade network model (Figure 3), it can be seen that the countries with larger weighted out degree in the current global trade network include Chile, Peru, Indonesia, Mexico, Australia, etc. They are the main exporting countries of global copper ore. The countries with larger weighted in

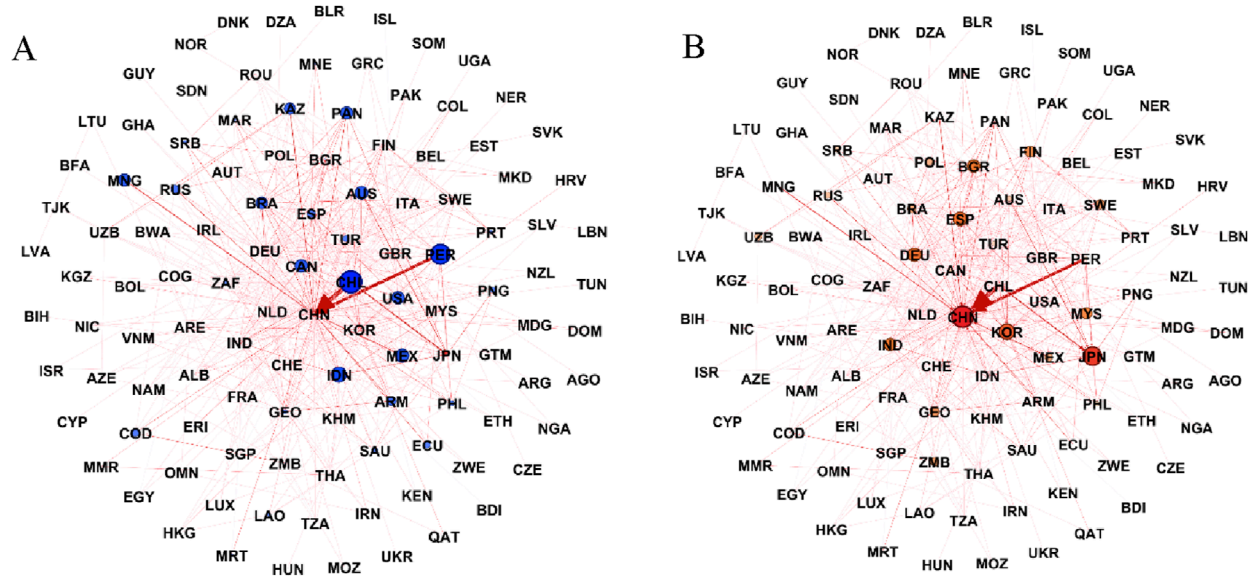


FIGURE 3
Global copper supply chain network model **(A)** weighted out degree, **(B)** weighted in degree).

degree include China, Japan, South Korea, etc., which are the main importing countries of global copper ore.

By counting the meso-centrality, proximity centrality and eigenvector centrality of each node, the important nodes in the global copper ore trade network are shown in the following table (Table 1). Apart from the major import and export countries, Netherlands, United Arab Emirates, Turkey, United Kingdom and other countries have strong control over the global trade network. Netherlands, Spain and Turkey are also important nodes in the network, they are typically interconnected with major trading hubs.

TABLE 1 Key nodes of the global copper ore trade network.

Importance indicators	Top 10 country names
Weighted indegree	China, Japan, Rep. of Korea, Spain, Germany, Bulgaria, India, Malaysia, Finland, Georgia
Weighted outdegree	Chile, Peru, Indonesia, Mexico, Australia, USA, Canada, Panama, Mongolia, Brazil
Weighted degree	China, Chile, Peru, Japan, Indonesia, Rep. of Korea, Spain, Mexico, Australia, Brazil
Betweenness centrality	China, Netherlands, USA, United Arab Emirates, Germany, Turkey, Chile, United Kingdom, Canada, Italy
Closeness centrality	Lithuania, Latvia, China, USA, Chile, Canada, Peru, Australia, Brazil, Mexico
Eigenvector centrality	China, Netherlands, Spain, Rep. of Korea, Chile, Germany, Malaysia, Turkey, Canada, Indonesia

3.2 Analysis of the historical evolutionary characteristics of the global copper ore trade network

This paper constructs a copper ore supply chain trade network from 1990 to 2021, and analyzes the evolutionary characteristics of the global copper ore trade pattern. From 1990 to 2021, the total volume of global copper ore trade continued to grow, with China's trade volume gradually dominating the world since 2000 (Figure 4). The total number of countries participating in the global copper ore trade network grew from 43 to 114. The number of trade connecting edges grew from 99 to 580 during the 30-year period. Both the number of nodes and connected edges reached a peak in 2015, then began to decline, and has been growing since 2020 (Figure 5). Combined with the global copper price fluctuations, the changes in the number of nodes and continuous edges are obvious hysteresis to copper price peak.

The topological structure characteristics of the network can be judged based on the average degree, average path length, average network density, average clustering coefficient, and degree off the number of ties of the network. The average degree of network, average density of network and average path length are often used to measure the closeness of trade network. The greater the average degree, the greater the density of network, and the smaller the average path length, the more closely connected the network structure is. The average clustering coefficient is used to measure the connectivity of trade between countries, the larger the average clustering coefficient, the better the trade connectivity between countries.

The average degree of the global copper ore trade network has gradually increased over time, with the average path length

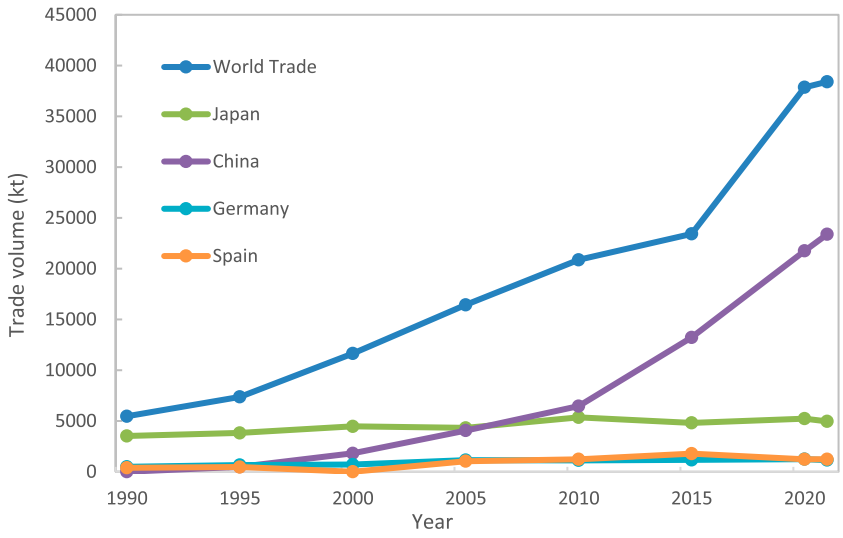


FIGURE 4
Graph of changes in global and major countries' copper ore trade volumes (1990-2021).

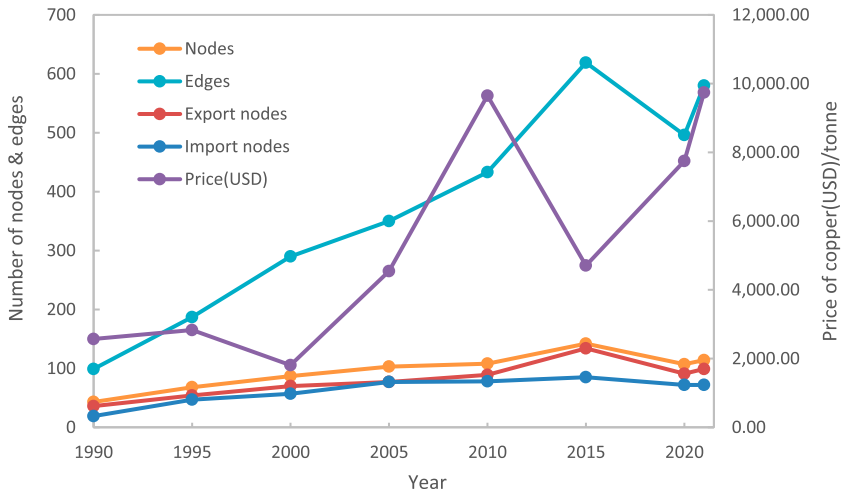


FIGURE 5
Graph of changes in the size of the global copper ore trading network (1990-2021).

between any two countries in the network ranging from 2.5 to 3. Since 1995, there has been an overall decreasing trend, indicating that the network tightness has been gradually increasing since 1995. In terms of the change in the average network density, it is highest in 1990, decreasing between 1990 and 2005, increasing from 2005 to 2010, decreasing from 2010 to 2015, reaching a minimum in 2015, and increasing after 2015. This may due to the fact that the network in 1990 contained fewer nodes and the network size was smaller, thus having a smaller average path length and a larger network density. the growth rates of the number of nodes from 1990 to 2020 were 58%, 28%, 18%, 4%, 31%, 25%, and 7%, respectively. The number of nodes increased significantly from 1990 to 2005 and 2015, the network size increased rapidly, and the average density of the network has decreased (Figure 6). Overall, the global copper ore trade network is gradually expanding, and the degree of network tightness is affected by the expansion rate. So, more low-degree nodes entering the network during the rapid expansion phase, resulting in a decrease in the average network density. There is an increase in network tightness during the slow expansion phase through the establishment of more connections between nodes.

The average clustering coefficient of the network keeps increasing, the small-world characteristic tends to be obvious (Figure 7). The trade connectivity among countries becomes stronger. Combined with the clustering coefficient-degree correlation distribution (Figure 8), it can be seen that the change of clustering coefficient with node degree is not obvious before 2010, and after 2010. The clustering coefficient decreases with the

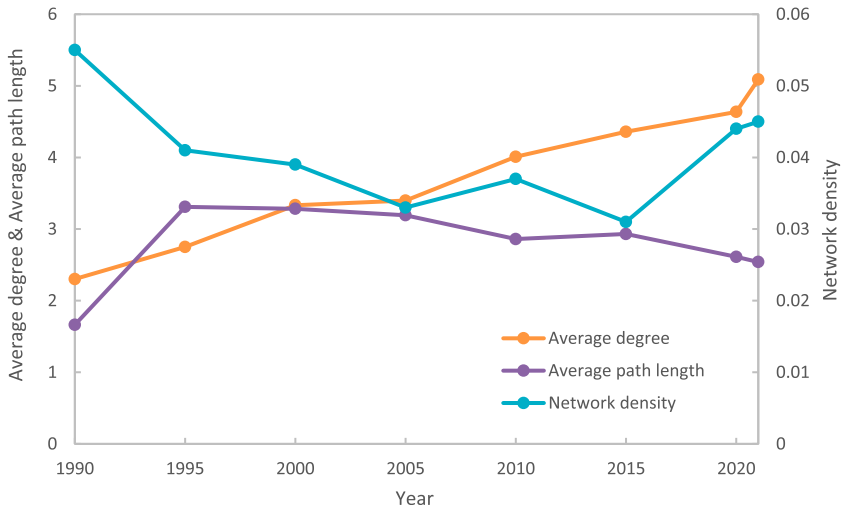


FIGURE 6
Characterization of the topology of the global copper ore trade network (1990-2021).

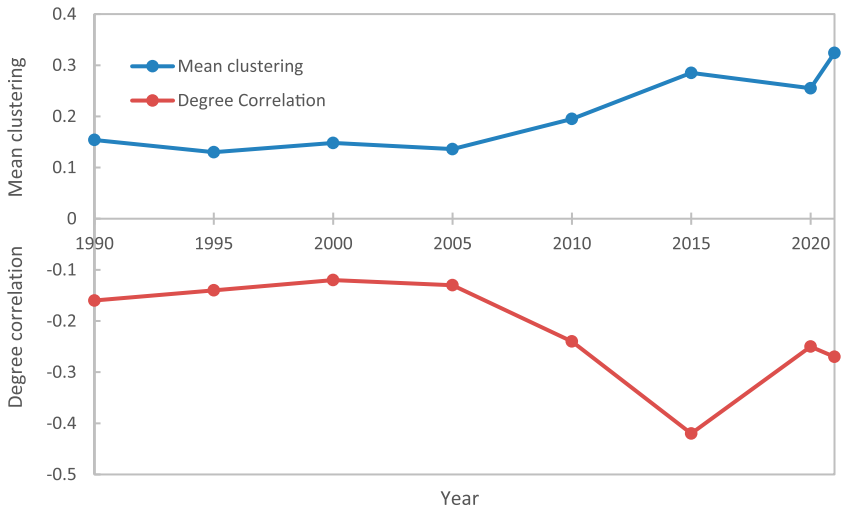


FIGURE 7
Graph of the change in clustering coefficient and degree correlation coefficient of global copper ore trade network (1990-2021).

increase of node degree, and the network shows the typical power-law distribution characteristics. There are hub nodes in the network, and the hub nodes make the distance between each other shorter by connecting a large number of low-degree nodes. It greatly shortens the path length and strengthens the connectivity of the network.

The degree correlation coefficients of the network are all negative, and the network as a whole shows the characteristics of a heterogeneous network. The nodes with large degrees in the network tend to establish connections with nodes with small degrees. The heterogeneity has weakened since 1990 to 2005, and the network tends to be neutral. From 2005 to 2015, the heterogeneity of the network strengthens significantly, reaching a peak in 2015 and weakening by 2020. This may due to the evolution of the network in which a small number of nodes gradually hold a large amount of trade resources. Exporting countries are restricted by the natural conditions of copper resource endowment, and importing countries are mainly concentrated in Central Asia and some countries in Europe. Especially since 2005, with the accelerated industrialization of China, the demand for copper has increased dramatically, highlighting the import demand and the need to expand import sources. The import targets tend to be low-degree nodes with fewer trade links, with more copper resource countries joining the global copper trade supply chain.

3.3 Simulation of the evolutionary process of the global copper ore trade network

According to the above spatial weighted complex network model based on the gravity model, the expected trade flows are calculated.

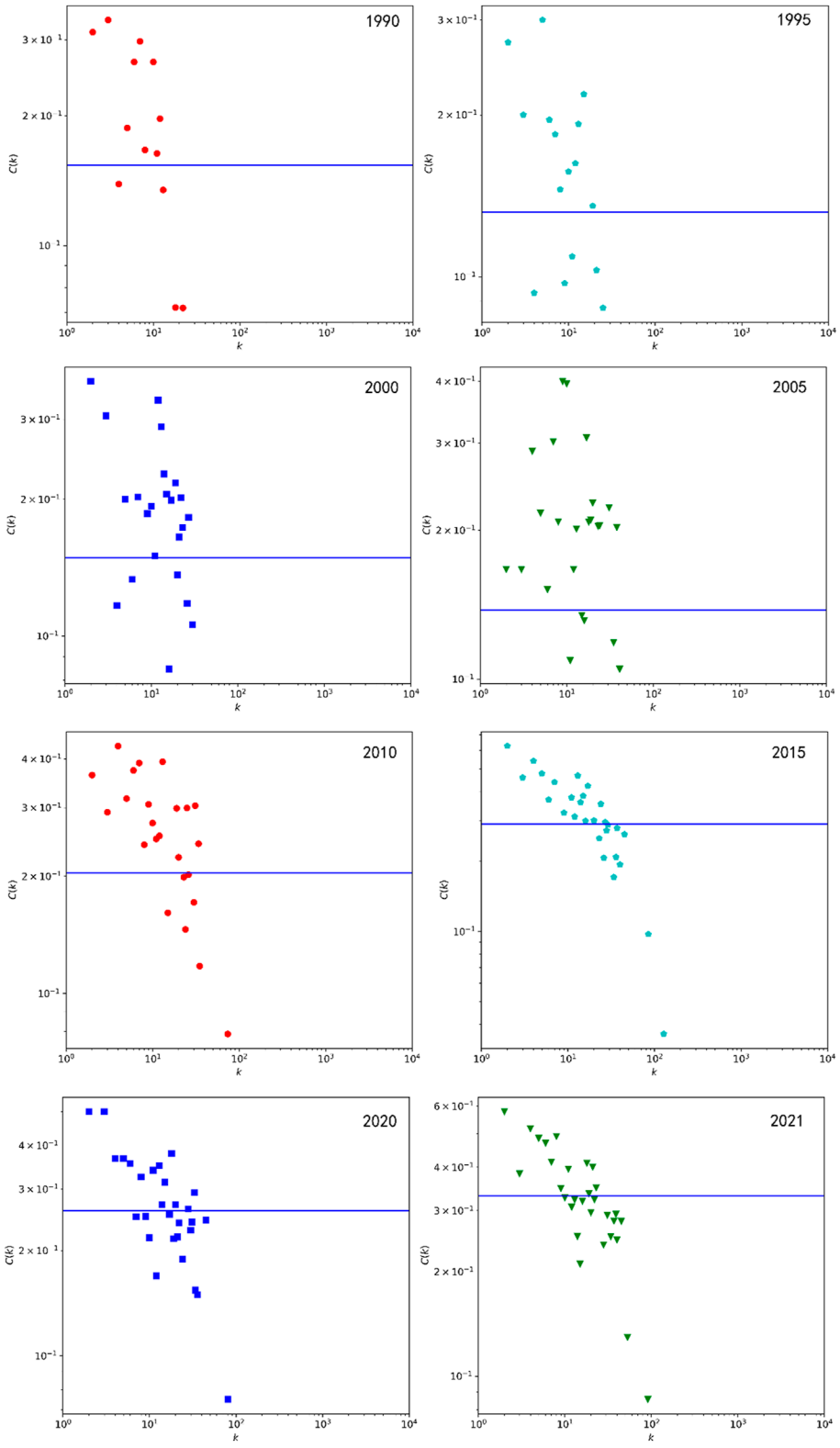


FIGURE 8
Global copper ore trade network aggregation coefficient - degree correlation distribution.

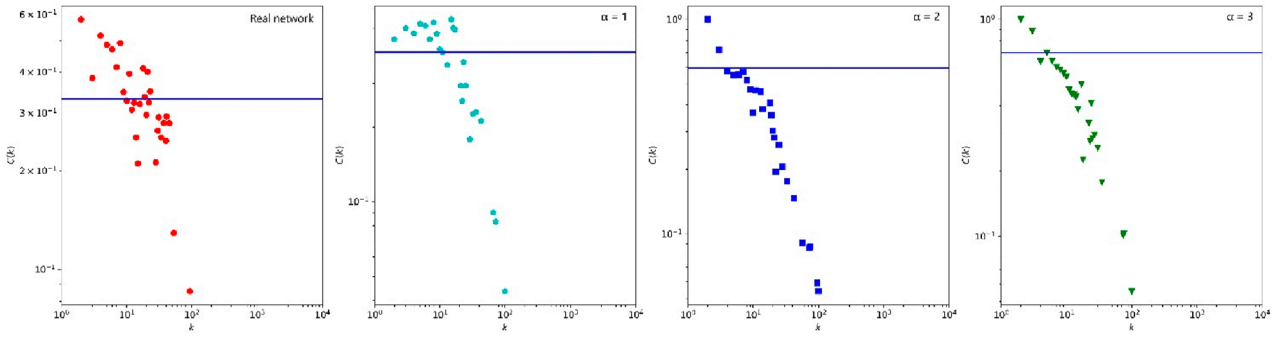


FIGURE 9
Simulation network clustering coefficient - degree correlation distribution (α = 1, 2, 3).

TABLE 2 Statistics of basic characteristics of simulated networks (α = 1, 2, 3).

	Average path length	Average degree	Network diameter	Network density	Average clustering coefficient	Modularity	(Pearson)
Real Networks	2.540	5.088	6	0.045	0.324	0.38	−0.258
α = 1	2.034	5.088	5	0.045	0.390	0.30	−0.336
α = 2	1.609	5.088	4	0.045	0.592	0	−0.626
α = 3	1.515	5.088	3	0.045	0.701	0	−0.667

It contains 114 countries as nodes in the global copper ore supply chain including 99 exporting countries and 72 importing countries. According to the size of the expected trade flows, the node pairs with larger expected flows are connected preferentially until the number of connected edges reaches 580 of the real network. Since there are duplicate countries among the 99 exporting countries and 72 importing countries, the network self-loop is not considered in the process of this modeling. So the self-loop is eliminated first in the model construction, and the connected edges are deferred until the number of connected edges reaches 580.

According to the gravity model equation, the values of α and ε depend on the dependence of the network on node adaptation (trade demand) and on geographical factors (transportation cost). They are the main parameters of the simulation experimental study in this paper. First, assuming ε = 1, α = 1, α = 2, α = 3 are set to experiment the dependence of the network on node adaptation degree. The clustering coefficient-degree correlation distribution of the model network shows a power-law distribution (Figure 9). As the value of α increases, the network topology exhibits a hierarchical structure. The network gives priority to connecting node pairs with large trade demand, and the hierarchical structure is strengthened.

The analysis of the structural characteristics of the simulated network can be seen (Table 2), with the increase of α value, the average degree and network density remain unchanged, the average path length decreases. It indicates that the network tends to be closely connected. The average clustering coefficient increases, and the trade connectivity of countries increases. The number of degree off links (Pearson) decreases significantly, indicating that the characteristics of the heterogeneous network are amplified. Network modularity is used to detect the community structure of the network. The higher the modularity, the better the structure of the divided community. If the modularity is 0 or negative, it means the whole network is a single community or each node is a separate community. Among them, the modularity degree when α = 2 and 3 is 0, and the community structure does not match the real network. Meanwhile, the average path length and clustering coefficient at α = 1 are closer to the real network, therefore, α = 1 is selected for the next analysis.

Specifically, when the parameters α and ε are both set to 1 Among them, ε = 0.5, ε = 1 when all nodes of the network are connected, and ε = 2 when some nodes of the network are not connected (Figure 10). As the value of ε increases and the geographic restriction strengthens, the network hierarchy gradually weakens (Figure 11). When the geographic factor strengthens, the network becomes homogeneous and the phenomenon of degree negative correlation gradually disappears, such as when ε = 2, the nearest neighbor average degree value is almost a constant, indicating that the network has no degree correlation (Figure 12).

Combined with the above comparison, the basic structural characteristics of the statistical simulation network (Table 3) show that as the value of ε increases, the average path length increases, the network diameter increases, the network density remains the same, the average clustering coefficient becomes smaller, the modularity becomes smaller, and the hierarchical and community characteristics of the network gradually disappear.

When ε is set to 1, the model network is closer to the real network, and the nodes with higher weighting degree include China, Chile, Peru Japan, and Indonesia. They are completely consistent

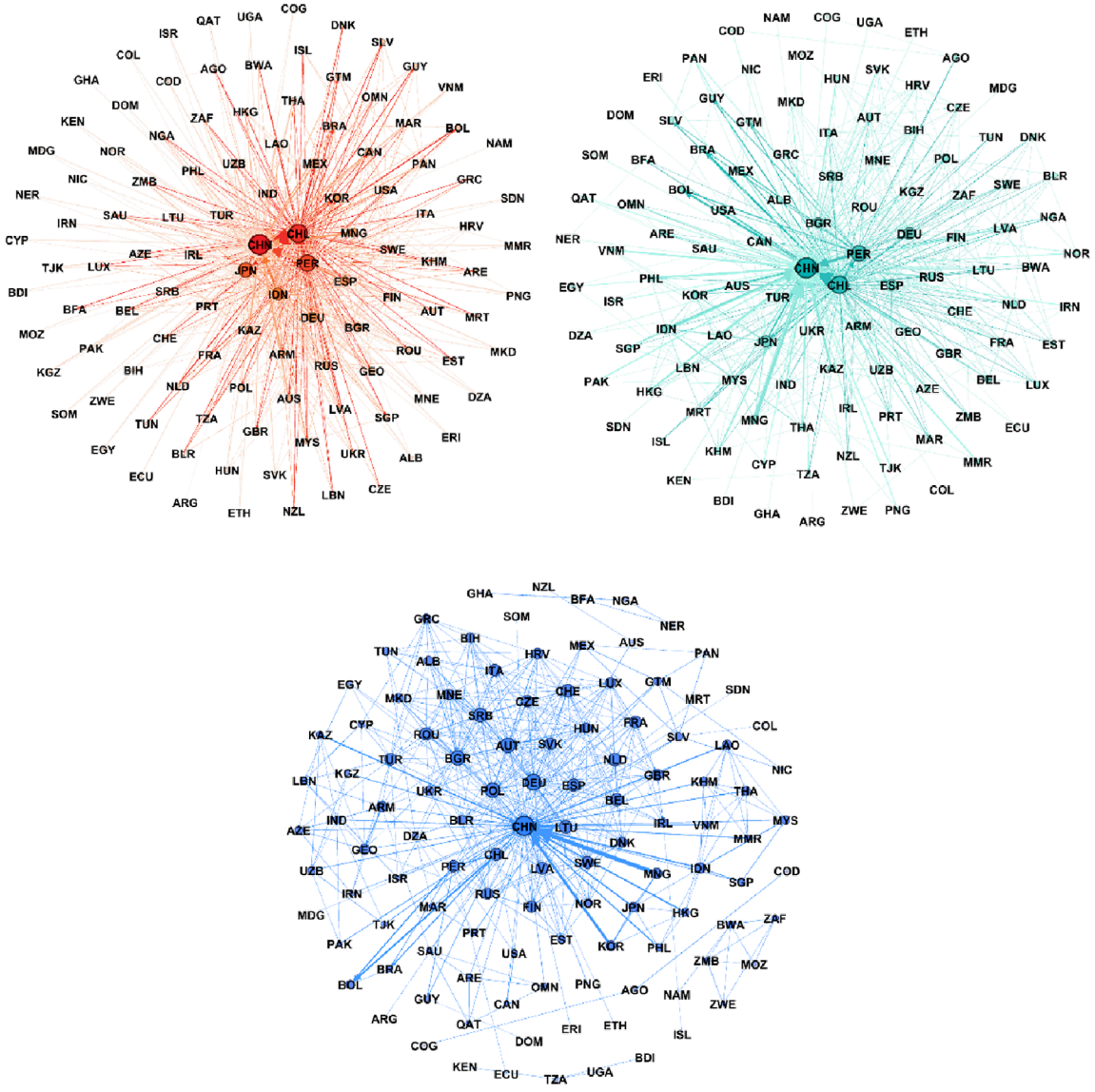


FIGURE 10
Global copper supply chain network model (ε = 0.5, 1, 2).

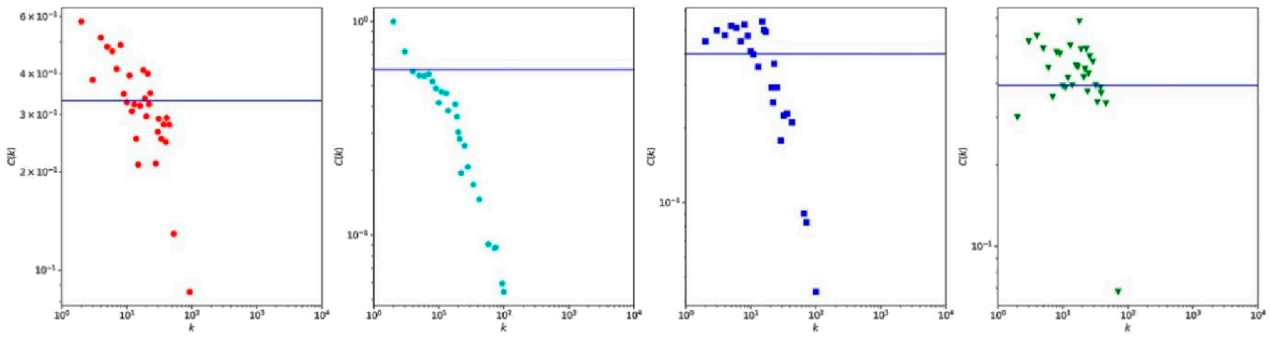


FIGURE 11
Simulation network clustering coefficient-degree correlation distribution (ε = 0.5, 1, 2).

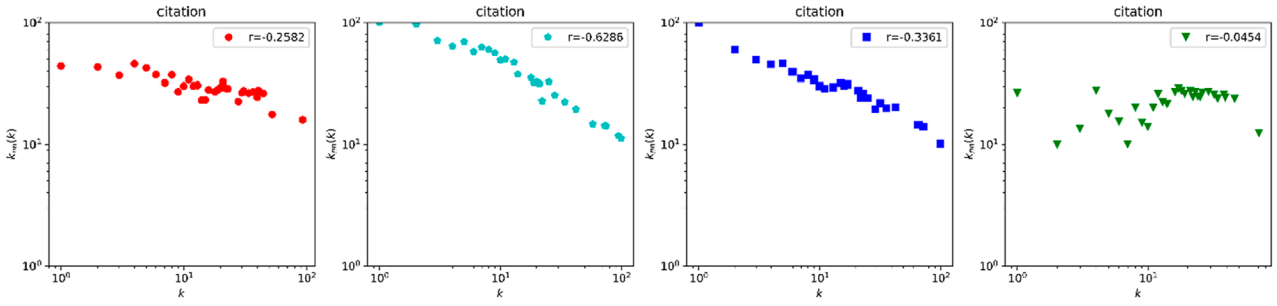


FIGURE 12
Simulation network degree - degree correlation distribution (ε = 0.5, 1, 2).

TABLE 3 Statistics of basic characteristics of simulated networks (ε = 0.5, 1, 2).

	Average path length	Average degree	Network diameter	Network density	Average clustering coefficient	Modularity	Degree correlation coefficient (Pearson)
Real Networks	2.540	5.088	6	0.045	0.324	0.38	−0.258
ε = 0.5	1.628	5.088	4	0.045	0.594	0.143	−0.628
ε = 1	2.034	5.088	5	0.045	0.390	0.30	−0.336
ε = 2	2.186	5.088	7	0.045	0.369	0.51	−0.045

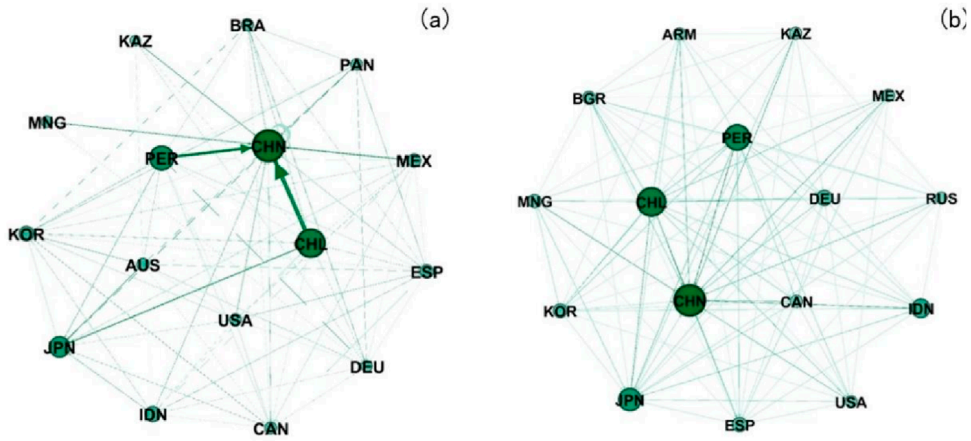


FIGURE 13
Global copper supply chain simulation network. [**(A)** original network, **(B)** simulation network].

with the real network. There are 17 of the top 20 countries in the weighted degree are consistent with the real network (Figure 13). It proves that the model can basically describe the driving pattern of establishing trade links between nodes. This model relates positively to the anticipated trade flow between nations and inversely to the intervening distance. It indicates that trade demand and the spatial distance are main driving forces for establishing trade ties among countries. Usually, trade ties are established with neighboring exporting countries that are relatively close in distance. With the increase in trade demand, the cost is limited, there will always be priority investment in the construction of the most urgent, the largest demand for trade flow. On this foundation, a scientific evaluation of the stability and potential risks associated with supply chain structures can be undertaken.

The model reflects the ideal structure of the global copper ore trade network. But the trade relationships in the real network are simultaneously influenced by many factors, such as national geopolitical relations, market prices, import and export policies,

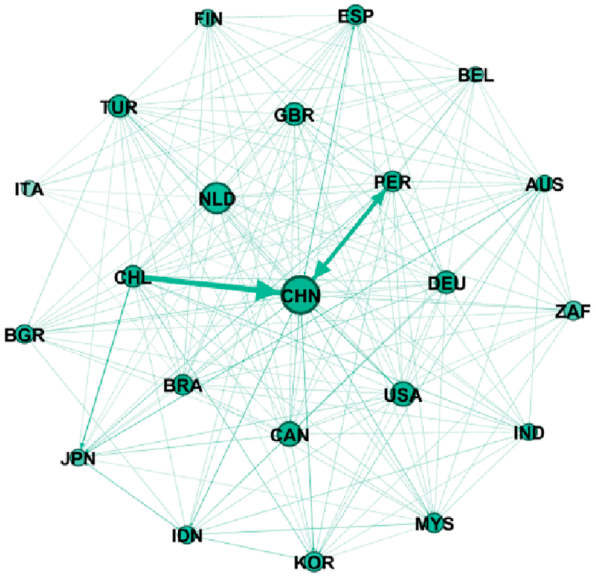


FIGURE 14
Global copper supply chain in 2022.

regional trade agreements and others. There are some different relationships from the simulation results in this paper. The nodes in the real network that have not yet established trade links deserve further focused attention.

3.4 Forecast of potential trading countries

In recent years, China has been the world's largest importer of copper ore. In this paper China is taken as an example to forecast the potential trading countries using the spatial weighted complex network model proposed above. Based on the data of 2021, the above model is used to simulate the evolution of the global copper supply chain structure and predict the changes of the global copper supply chain structure in 2022. The simulation results are compared with the real network to evaluate the accuracy of the model.

In 2021, China imported 23 million tons of copper ore from 59 countries. The major exporting countries are Chile, Peru, Kazakhstan, Mongolia, Mexico, Indonesia. These countries above account for nearly 80% of China's copper imports. The spatial weighted complex network model was used with $\alpha = 1$ and $\varepsilon = 1$. According to the simulation results, in 2022, there was little change in the major exporting countries, such as Chile, Peru, Kazakhstan, Mongolia, Mexico, Indonesia (Figure 14). And these countries above account for nearly 80% of China's copper imports (Figure 15).

From the perspective of network structure, China's copper imports come from more diverse sources. The number of countries trading copper with China has grown to 83 in 2022. These include 80 countries in the list of countries simulated, the percentage is as high as 96%. This can also show that with the change of the global geographical situation, China is seeking new trading partners, expand the diversity of copper import sources, and more countries have joined the copper trade with China.

4 Conclusion

Based on the global copper ore trade import and export data from 1990 to 2022, this paper constructs a global copper ore trade network model using complex network theory. The historical evolution characteristics and structural change patterns of the global copper ore trade network are analyzed. The evolution process of the global copper ore trade network structure is simulated using a spatially weighted complex network model. Based on these findings, we propose an evolution pattern for the global copper ore trade network under the joint influence of national trade demand and transportation costs. The main conclusions drawn in this paper are as follows:

(1) The spatial heterogeneity in the distribution of mineral resources among exporting nations, coupled with the economic progress of importing countries, exerts a profound influence on the evolution of the global trade network structure. Major importing economies, notably Japan, China, Germany, and Spain, primarily procure copper resources from significant trade partners such as Chile and Canada. Concurrently, these importing nations are diversifying their import bases to encompass resource-rich countries situated in closer geographic proximity, thereby enhancing their supply chain resilience and strategic sourcing capabilities within the global trade framework.

(2) Since 1990, the total volume of global copper ore trade has exhibited a sustained growth trajectory, accompanied by a gradual expansion of the global copper ore trade network. The degree of network closeness exhibits fluctuations that are contingent upon the rate of this expansion. Notably, the small-world and heterogeneous attributes of the network have become increasingly apparent, underscoring an intensification of trade connectivity among nations.

(3) The evolutionary pattern of the global copper ore trade network can be accurately characterized using the spatial weighted complex network evolution model. This model relates positively to the anticipated trade flow between nations and inversely to the intervening distance. When both parameters α and ε are set to 1, the structure of the simulated copper ore trade network aligns more closely with the real network. On this foundation, a scientific evaluation of the stability and potential risks associated with supply chain structures can be undertaken.

Building upon the insights gained from our preceding analysis, we undertake a comprehensive examination of the structural attributes and evolutionary forces shaping the global copper supply chain. Subsequently, we offer a suite of strategic recommendations targeted at optimizing the international copper trade dynamics at a comprehensive level. Presently, the exportation of global copper ore is predominantly concentrated in key regions such as

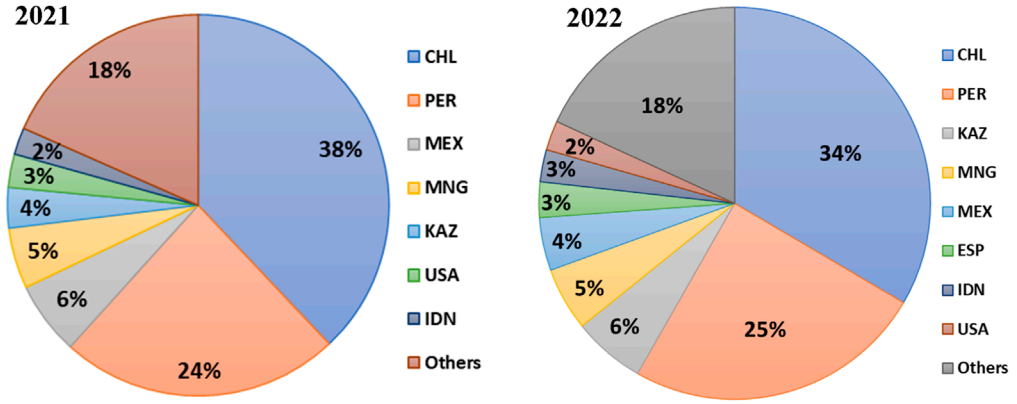


FIGURE 15
Source of China's copper imports in 2021 and 2022.

Chile, Peru, Indonesia, Mexico, and Australia. The potential trade relationships projected from this concentration merit particular scrutiny. It is advisable for importing countries such as China, Japan, Rep. of Korea, Spain, Germany to diversify copper sources and securing a stable supply chain. And suggestions for export countries such as Chile, Peru, Indonesia, Mexico, Australia to enhance production resilience and maintain trade advantages. For major trade intermediary countries such as the Netherlands, United Arab Emirates, Turkey, it is essential to stabilize the trade partnerships established with other nodes and actively expand their coverage and influence within the network. Furthermore, establishing copper trade relations with neighboring countries is recommended as a strategic approach to ensure the stability and resilience of the supply chain.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

QC: Conceptualization, Formal Analysis, Funding acquisition, Project administration, Writing–original draft, Writing–review and editing. KW: Conceptualization, Formal Analysis, Methodology, Writing–original draft, Writing–review and editing. YZ: Methodology, Software, Writing–review and editing. QG: Methodology, Software, Writing–original draft. JX: Data curation, Visualization, Writing–original draft. TL: Data curation, Visualization, Writing–original draft. GZ: Supervision, Writing–review and editing. QL: Project administration, Supervision, Writing–review and editing. ZL: Project administration, Supervision, Writing–review and editing. XR: Data curation, Writing–original draft. CS: Data curation, Writing–original draft. YD: Methodology, Software, Writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by National Natural Science Foundation of China (NSFC) [grant numbers 42271281]; Deep Earth Probe and Mineral Resources Exploration - National Science and Technology Major Project [2024ZD1002006]; China Geological Survey, Ministry of Natural Resources [grant number DD20230040; DD20221694; DD20211405]; and Chinese Academy of Engineering strategic research and consulting project [grant numbers 52922023002].

Acknowledgments

The authors thank the reviewers for helpful comments on the draft manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Kumar BH, Reddy NR, Kumar MCS. Cu-rich copper indium sulfide thin films deposited by co-evaporation for photovoltaic applications. *J Mater Sci Mater Electronics* (2023) 34:341. doi:10.1007/s10854-022-09441-w

2. Abbas S, Saqib N, Shahzad U. Global export flow of Chilean copper: the role of environmental innovation and renewable energy transition. *Geosci Front* (2024) 15:101697. doi:10.1016/j.gsf.2023.101697

3. Tang J, Yang H, Chen H, Li F, Wang C, Liu Q, et al. Potential and future direction for copper resource exploration in China. *Green Smart Mining Eng* (2024) 1:127–31. doi:10.1016/j.gsme.2024.06.001

4. Rivera N, Guzmán J, Jara J, Lagos G. Evaluation of econometric models of secondary refined copper supply. *Resour Pol* (2021) 73:102170. doi:10.1016/j.resourpol.2021.102170

5. Alexander D, Merkert R. Applications of gravity models to evaluate and forecast US international air freight markets post-GFC. *Transport Policy* (2021) 104:52–62. doi:10.1016/j.tranpol.2020.04.004

6. Guj P, Schodde R. Will future copper resources and supply be adequate to meet the net zero emission goal? *Geosystems and Geoenvironment* (2025) 4:100320. doi:10.1016/j.geogeo.2024.100320

7. Kang X, Wang M, Wang T, Luo F, Lin J, Li X. Trade trends and competition intensity of international copper flow based on complex network: from the perspective of industry chain. *Resour Pol* (2022) 79:103060. doi:10.1016/j.resourpol.2022.103060

8. Guan Q, An H. The exploration on the trade preferences of cooperation partners in four energy commodities' international trade: crude oil, coal, natural gas and Photovoltaic. *Appl Energ* (2017) 203:154–63. doi:10.1016/j.apenergy.2017.06.026

9. Chen Z, An H, An F, Guan Q, Hao X. Structural risk evaluation of global gas trade by a network-based dynamics simulation model. *Energy* (2018) 159:457–71. doi:10.1016/j.energy.2018.06.166

10. Nassar N, Brainard J, Gulley A, Manley R, Matos G, Lederer G, et al. Evaluating the mineral commodity supply risk of the U. S. manufacturing sector. *Sci Adv* (2020) 6:8647. doi:10.1126/sciadv.aay8647

11. Althaf S, Babbitt C. Disruption risks to material supply chains in the electronics sector. *Resour Conservation and Recycling* (2020) 167:105248. doi:10.1016/j.resconrec.2020.105248

12. Ryter J, Fu X, Bhuwalka K, Roth R, Olivetti EA. Emission impacts of China's solid waste import ban and COVID-19 in the copper supply chain. *Nat Commun* (2021) 12:3753. doi:10.1038/s41467-021-23874-7

13. Brink S, Kleijn R, Sprecher B, Tukker A. Identifying supply risks by mapping the cobalt supply chain. *Resour Conservation and Recycling* (2020) 156:104743. doi:10.1016/j.resconrec.2020.104743

14. Pizarro J, Sainsbury B, Hodgkinson J. Adaptation options assessment for the Australian uranium supply chain focused on the Olympic Dam and Ranger Mines. *Environ Development* (2020) 37:100610. doi:10.1016/j.envdev.2020.100610

15. Mou M, Zhou Y, Zheng W, Xie Y. Integration and modeling of multi-energy network based on energy hub. *Complexity* (2022) 2022:2698226. doi:10.1155/2022/2698226

16. Schnebele E, Jaiswal K, Luco N, Nassar N. Natural hazards and mineral commodity supply: quantifying risk of earthquake disruption to South American copper supply. *Resour Pol* (2019) 63:101430. doi:10.1016/j.resourpol.2019.101430

17. Zhang H, Xi W, Ji Q, Zhang Q. Exploring the driving factors of global LNG trade flows using gravity modelling. *J Clean Prod* (2018) 172:508–15. doi:10.1016/j.jclepro.2017.10.244

18. Shen J, Huang S. Copper cross-market volatility transition based on a coupled hidden Markov model and the complex network method. *Resour Pol* (2022) 75:102518. doi:10.1016/j.resourpol.2021.102518

19. Ye H, Li Z, Li G, Liu Y. Topology analysis of natural gas pipeline networks based on complex network theory. *Energies* (2022) 15:3864. doi:10.3390/en15113864

20. Hao M, Tang L, Wang P, Wang H, Wang Q, Dai T, et al. Mapping China's copper cycle from 1950-2015: role of international trade and secondary resources. *Resour Conservation Recycling* (2023) 188:106700. doi:10.1016/j.resconrec.2022.106700

21. Yuan ZW, Ji HR, Zhang L, Liu L. Sustaining the United States' copper resource supply. *J Clean Prod* (2023) 429:139636. doi:10.1016/j.jclepro.2023.139636

22. Zhang W, Ye Y, Li Z, Xian J, Wang T, Liu D, et al. The coupled awareness-epidemic dynamics with individualized self-initiated awareness in multiplex networks. *Front Phys* (2024) 12:1437341. doi:10.3389/fphy.2024.1437341

23. Suárez V, Montoya-Torres J, Sepúlveda-Rojas J. A gravitational model extended by institutional and cultural factors for Colombian foreign trade. *Management Sci Lett* (2021) 11:2313–22. doi:10.5267/J.MSL.2021.6.003

24. Duenas M, Fagiolo G. Global trade imbalances: a network approach. *Advcomplex Syst* (2014) 17:1450014. doi:10.1142/S0219525914500143

25. Zhang YT, Zhou WX. Structural evolution of international crop trade networks. *Front Phys* (2022) 10:926764. doi:10.3389/fphy.2022.926764

26. Liu M, Guo Q, Liu J. Effect of network structure on the accuracy of resilience dimension reduction. *Front Phys* (2024) 12:1420556. doi:10.3389/fphy.2024.1420556

27. Li B, Li H, Ren S, Liu H, Wang G. Commodity supply risk assessment of China's copper industrial chain: the perspective of trade network. *Resour Pol* (2023) 81:103297. doi:10.1016/j.resourpol.2023.103297

28. Barabási AL. *Network science*. Cambridge University Press (2016).

29. Zhang L, Jian X, Ma Y. Analysis of differences in fossil fuel consumption in the world based on the fractal time series and complex network. *Front Phys* (2024) 12:1457287. doi:10.3389/fphy.2024.1457287

30. Ma Y, Wang MX, Li X. Analysis of the characteristics and stability of the global complex nickel ore trade network. *Resour Pol* (2022) 79:103089. doi:10.1016/j.resourpol.2022.103089

31. Xi X, Zhou J, Gao X, Liu D, Zheng H, Sun Q. Impact of changes in crude oil trade network patterns on national economy. *Energy Econ* (2019) 84:104490. doi:10.1016/j.eneco.2019.104490

32. Aydın U, ülengin B. Analyzing air cargo flows of Turkish domestic routes: a comparative analysis of gravity models. *J Air Transport Management* (2022) 102:102217. doi:10.1016/j.jairtraman.2022.102217

33. Chenaf-Nicet D, Rougier E. The effect of macroeconomic instability on FDI flows: a gravity estimation of the impact of regional integration in the case of Euro-Mediterranean agreements. *Int Econ* (2016) 145:66–91. doi:10.1016/j.inteco.2015.10.002

34. Kwon O, Jung W. Intercity express bus flow in Korea and its network analysis. *Physica A: Stat Mech its Appl* (2012) 391:4261–5. doi:10.1016/j.physa.2012.03.031

35. Lundmark R. Analysis and projection of global iron ore trade: a panel data gravity model approach. *Miner Econ* (2018) 31:191–202. doi:10.1007/s13563-017-0125-8

36. Akerman A, Leuven E, Mogstad M. Information frictions, internet, and the relationship between distance and trade. *Am Econ J Appl Econ* (2022) 14:133–63. doi:10.1257/app.20190589

37. Hsieh TH, Zhu YB, Huang KL. Using the gravitational mixed models to analyze the impact of China's foreign direct investment along with the Belt and Road countries on trade flows. *Front Psychol* (2022) 13:960722. doi:10.3389/fpsyg.2022.960722

38. Zhao N, Liu Q, Wang H, Yang SL, Li PZ, Wang J. Estimating the relative importance of nodes in complex networks based on network embedding and gravity model. *J King Saud Univ - Computer Inf Sci* (2023) 35:101758. doi:10.1016/j.jksuci.2023.101758

39. Dabbaghfeizi S, Naderan A, Tavakoli-Kashani A. Aggregate crash prediction model based on gravity model: introducing crash risk distribution concept. *Transportation Eng* (2025) 19:100297. doi:10.1016/j.treng.2024.100297

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 14 March 2025
DOI 10.3389/fphy.2025.1557361

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Yuan Ni,
Beijing Information Science and Technology University, China
Sui Guo,
Jiangsu Normal University, China

*CORRESPONDENCE
Rui Kong,
kongrui@cugb.edu.cn

RECEIVED 08 January 2025
ACCEPTED 10 February 2025
PUBLISHED 14 March 2025

CITATION
Wang M, Kong R, Luo J and Hao W (2025) Co-movement forecasting between consumer sentiment and stock price in e-commerce platforms using complex network and entropy optimization. *Front. Phys.* 13:1557361. doi: 10.3389/fphy.2025.1557361

COPYRIGHT
© 2025 Wang, Kong, Luo and Hao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Co-movement forecasting between consumer sentiment and stock price in e-commerce platforms using complex network and entropy optimization

Mingyue Wang[1], Rui Kong[1,2]*, Jianfu Luo[3] and Wenjing Hao[1]

[1]School of Economics and Management, China University of Geosciences, Beijing, China, [2]Key Laboratory on Resources and Environment Capacity under Ministry of Land and Resources of People's Republic of China, Beijing, China, [3]School of Engineering, Hulunbuir University, Hailar, Inner Mongolia, China

Stock price and consumer sentiment consistently serve as pivotal economic indicators for the performance and growth of e-commerce enterprises. It is essential to comprehend and forecast the co-movement between the two to inform financing and investment decision-making effectively. Prior research has focused on predicting individual indicators, but not much of them attempt to forecast their co-movement. We propose a novel Rule Combination based on Bivariate Co-movement Network (RC-BCN) approach for bivariate co-movement forecasting. Bivariate co-movement features extracted utilizing the BCN's topological nature instruct the entropy optimization in order to enhance the RC-BCN's predictions. We conduct four sets of experiments on 1,135 data sets from JD.com between 2018 and 2022, where consumer sentiment is measured using text sentiment analysis of online reviews. The results indicate that RC-BCN's prediction accuracy reaches at most 91% under distortion preference and is improved by 18% compared without entropy optimization. This study highlights the value of complex network and entropy theory in forecasting bivariate co-movement for e-commerce enterprises.

KEYWORDS

consumer sentiment, stock price, bivariate co-movement network, entropy optimization, text sentiment analysis

Highlights

- Devising a novel approach called Rule Combination based on Bivariate Co-movement Network (RC-BCN) to solve co-movement prediction problem.
- Predicting the co-movement between Consumer Sentiment (CS) and Stock Price (SP) for e-commerce companies.
- Bi-objective optimization achieved through entropy optimization based on BCN topology.
- RC-BCN outperforms in co-movement between CS and SP forecasting.

- Findings in this study have implications for listed e-commerce companies financing decisions and investor decision-making.

1 Introduction

For publicly listed enterprises, particularly those in the customer-centric retail sector, forecasting the co-movement between Consumer Sentiment (CS) and Stock Price (SP) is crucial for guiding corporate financing decisions and providing valuable insights for investor decision-making. E-commerce enterprises, as typical customer-centric businesses, are significantly influenced by CS in terms of their firm value. Additionally, the e-commerce industry has become a vital component of the global economy. CS and SP are economic signal towers of e-commerce platform performance. Numerous studies have independently examined and forecasted them [1–3]. In contrast, a few studies have integrated these two aspects and established a confirmed correlation between them [4, 5]. These studies have primarily explored the relationship between CS and SP via regression analysis, and employed machine learning to predict SP using CS, thereby enhancing predictive accuracy. Despite these advancements, the dynamic interdependence between CS and SP, particularly regarding co-movement forecasting, has received limited attention in existing literature.

Extant literature indicates that CS impacts SP through multiple pathways. (1) CS can indirectly affect investor behavior through media reports, social media interactions, and other channels, thereby impacting SP [6, 7]. (2) CS reflects the market preferences for a firm's products or services, signaling sales expectations and investor confidence, which in turn affect SP [8, 9]. (3) Some studies suggest that CS can directly and positively influence SP [4, 10]. Positive sentiment boosts market expectations of a firm's future performance, enhancing investor confidence, while negative sentiment undermines market trust, potentially leading to a decline in SP. In addition, related studies have found that: (1) there is a bidirectional causal relationship between sentiment and SP, with potentially positive or negative dynamic [11]. (2) the relationship between CS and SP is complex and nonlinear [5]; and (3) sentiment has a short-term transient effect on SP, suggesting that sentiment changes can serve as early signals of SP movements [11–13].

The relationship between CS and SP is dynamic. However, in contrast to previous studies focusing on the static aspect, this paper examines their dynamic interaction and co-movement forecasting, providing valuable insights into the dynamic relationship between consumer behavior and financial markets. Furthermore, investors can better capture market trends and mitigate potential investment risks by analyzing changes in CS. For customer-centric listed companies, understanding these dynamics enables them to improve financing efficiency and secure greater capital support. Hence, the research question in this study is: how to explore and forecast the co-movement of CS and SP? This research aims to uncover the co-movement's features and forecast them, serving as a reference for corporate financing decisions and investor investment strategies.

Forecasting the co-movement of CS and SP involves obtaining daily CS data and addressing the challenges of nonlinear time series forecasting. Currently, vast quantities of CORs are archived on e-commerce platforms, providing real-time insights into CS. Moreover, text sentiment analysis is employed to extract the CS from CORs at a finer time scale. Complex network theory has emerged as a powerful tool for analyzing nonlinear time series [14]. Given the dynamic, nonlinear relationship between CS and SP, and the complexity of their interaction patterns, complex network methods are particularly well-suited to uncover and forecast their co-movement. It excels in accurately capturing temporal fluctuations, leading to highly accurate predictions.

In response to the above problems, we propose a novel Rule Combination approach, integrating Bivariate Co-movement Network (RC-BCN) analysis with entropy optimization. It consists of four sequential steps: CS data acquisition, Bivariate Co-movement Network (BCN) construction, feature extraction, and co-movement prediction based on rule combination and entropy optimization. CS data acquisition uses web crawler and text sentiment analysis. We project bivariate time series data into phase space to construct BCN. Particularly, features derived from network topology are used to instruct entropy optimization innovatively. To validate the proposed approach, we conduct four sets of experiments on one dataset. The experiments show that the proposed approach outperforms baseline model.

Our work contributions are summarized as follows:

- Employ complex network to the problem of bivariate co-movement forecasting and develop a novel RC-BCN approach based on entropy optimization for the first time, which performs well.
- Propose a new feature extraction framework, using network topological nature to extract features by projecting bivariate time series into phase space.
- Design a fusion method of combining co-movement features with prediction methods through entropy optimization, which results in high prediction accuracy.

The rest of this paper is organized as follows. Section 2 discusses the related works. Section 3 describes the methodology. Section 4 presents the experiments, and Section 5 discusses the results. Section 6 concludes and outlooks the work.

2 Related works

To address the research question, we conducted a literature review from three aspects. Lastly, we highlighted the extant research gaps and outlined the objectives of our study.

2.1 Evaluating CS through online reviews

CS data can be sourced from two main categories: (1) Authoritative databases, which provide established metrics like the Consumer Confidence Index (CCI), and (2) Evaluating CS through online reviews. Authoritative databases may have limitations in providing real-time, flexible, and fine-grained analysis when supporting detailed research perspectives. Therefore, obtaining CS through CORS have become popular.

Evaluating CS through CORs primarily entails identifying the influencing factors and calculating CS. Various techniques,

such as text regression analysis, machine learning, and the Latent Dirichlet allocation, are employed to ascertain the factors and their respective weights that impact CS, consequently pinpointing the key determinants [15–17]. Additionally, evaluating CS through CORs predominantly utilizes text sentiment analysis methods based on dictionaries, machine learning, and deep learning [18–20]. Among these, the dictionary-based text sentiment analysis method is widely adopted, due to its simplicity. For instance, one approach involves constructing a CS evaluation system by content analysis, followed by the utilization of Probabilistic Linguistic Term sets and improved Analytic Hierarchy Process method to calculate CS based on word frequency statistics [21]. Other approaches include the Center Term Short Sentence Sentiment Orientation, the Probabilistic Linguistic Group Decision-FlowSort, and the Data Envelopment Analysis method to calculate sentiment intensity and dynamically measure CS across multiple dimensions [22–24].

These studies highlight the significance of CORs as a valuable data source for measuring CS, using relatively established evaluation methods. However, it is worth noting that CS in our study refers to consumers' overall sentiment on shopping platforms. While influenced by multiple factors, CS cannot be directly quantified as the sum of sentiment values derived from these individual factors. Hence, in quantifying CORs, we adopt a dictionary-based algorithm that matches overall CS vocabulary to measure the sentiment intensity of CORs. This approach effectively captures overall CS.

2.2 Bivariate co-movement features

Co-movement analysis between CS and SP typically focuses on their correlation, particularly in predicting SP volatility utilizing CS. However, limited research has been conducted on the co-movement features between CS and SP, while an extensive body of literature exists on bivariate co-movement features, showcasing more mature methodologies. The research subjects include inter-country currency exchange rates, the intersection of artificial intelligence and energy sectors, as well as the relationship between dirty and clean energy [25–27]. These research covers both static co-movement features, such as bivariate interdependence and quantile co-movement, and dynamic features, including co-movement structure and fluctuations. Although studies on CS and SP co-movement are typically classified as static, their co-movement is inherently dynamic.

Common methods for analyzing dynamic co-movement features include wavelet analysis, DCC-GARCH model, the Thick Pen method, Temporal Convolutional Network (TCN), and complex network analysis [25, 28–31]. Among these, wavelet analysis has high sensitivity to the selected wavelet basis function and scale parameters. The DCC-GARCH model generally assumes that the data are normally distributed. The Thick Pen method tends to be subjective in parameter setting. TCN depends on the quality of input data and is sensitive to the hyperparameter. In contrast, the co-movement between CS and SP does not necessarily conform to the normal distribution. Complex network analysis provides an objective means to visualize nonlinear and nonintuitive relationships between bivariate.

Scholars have employed symbolic dynamics methods to map time series data into BCN, and have utilized network indicators to analyze bivariate dynamic co-movement features. This approach is grounded on coarse-graining partitions of phase space, involves constructing coarse-grained rules to transform time series into symbols, defining symbol sequence vectors as nodes, establishing a network resulting from the transition frequency of the symbol sequence vectors in the time series [14]. The co-movement features are mainly extracted by using network topology indicators such as node strength, node centrality, weighted cluster coefficient and community detection [32, 33]. Despite employing various network indicators in prior research, there remains a gap in the literature regarding a comprehensive analytical framework for identifying bivariate co-movement features and guiding the development of a bivariate co-movement prediction model.

2.3 Bivariate co-movement prediction

The research area of bivariate co-movement forecasting includes different stock prices, the multi-step network traffic states, innovative and traditional financial assets in different regions [34–36]. Research topics in co-movement forecasting focus on the future trends of bivariate co-movement, including correlations between stock prices [34], crude oil prices volatility [37], and volatility spillovers across economic sectors [38]. Given the high volatility of both CS and SP, this study aims to forecast the co-movement trend between CS and SP, consistent with prior research.

In the absence of distinct regularities, nonlinear systems commonly adopt integrated forecasting models, incorporating techniques such as data mining, cross entropy, Convolutional Neural Network-Long Short Term Memory (CNN-LSTM), Back Propagation (BP) Neural Networks, and link prediction methods, to improve predictive accuracy [39–42]. The aforementioned methods, particularly machine learning techniques, have demonstrated significant efficacy in predicting bivariate co-movement by optimizing model hyperparameters. However, these methods face challenges in explaining the underlying mechanisms driving the observed results. In contrast, the complex network approach not only uncovers the co-movement characteristics of bivariate relationships but also effectively predicts their dynamics, eliminating the need for resource-intensive and time-consuming experimental processes.

Time series forecasting algorithms based on complex network have shown significant advancements in recent years. These methods, such as maximum node weight and maximum visibility forecasting have been developed based on the principles of complex network link prediction [43, 44]. The current widely used link prediction algorithms are often based on various node similarities, including local similarity, weighted similarity and random walk similarity [45–47]. Typically, the greater the similarity between two nodes in the network, the higher the likelihood of connectivity between them. While numerous algorithms exist for bivariate co-movement prediction, each validated in different fields, it is worth mentioning that the co-movement features within a network can vary significantly across different time series networks. As a result, predictions made by a single forecasting method may exhibit biases when compared to actual outcomes.

2.4 Research gap and objectives

Bivariate co-movement forecasting has been studied on mostly macroeconomic indicators. Limited by the availability of data, previous studies have less often landed on firm-level economic indicators, especially on the co-movement of CS and SP. For e-commerce companies, the abundance of CORs on e-commerce platforms provides valuable opportunities for in-depth analysis. Previous studies have developed various methods to extract co-movement features and demonstrated promising results in bivariate co-movement prediction. However, challenges persist in systematically extracting these features, applying them for prediction, and interpreting results, particularly in nonlinear systems. These gaps underscore the necessity of exploring novel approaches to predict the co-movement between CS and SP in the context of e-commerce enterprises.

Addressing the gaps identified in the extant literature, this study introduces a novel approach, grounded in complex network theory, to forecast the co-movement dynamics between CS and SP of e-commerce platforms. The basic approach is: design an algorithm for collecting and quantifying CORs on e-commerce platforms to obtain a more finely-grained CS; and then transform the time series data of CS and SP into BCN and extract the co-movement features based on the network's topology; construct a rule combinations model based on the co-movement features to forecast future co-movement mode. Moreover, model optimization involves entropy optimization guided by co-movement features.

3 Research methodology

3.1 Time-series data collection

The closing price is a crucial indicator that reflects daily market trends and is easily accessible through stock trading platforms. The process of obtaining the CS involved four key steps (see Figure 1). Firstly, CORs are collected using web crawling techniques. Secondly, a dictionary-based text sentiment analysis method is applied, utilizing the Jieba text segmentation tool to segment, analyze, and match the text with the dictionary for initial data processing. Then, sentiment values are calculated based on the part-of-speech and weights assigned by the dictionary. Finally, the obtained sentiment values that passed the test are used as the CS. The test algorithm demonstrated an accuracy of 84.77% when applied to the Tan Songbo hotel review corpus, confirming its practical applicability.

Since reviews consist of sentences, Equation 1 calculates sentiment scores at the sentence level, while Equation 2 aggregates them to determine the review-level sentiment score, and Equation 3 standardizes it.

$$\mathbf{Sentence'sscore} = \left(\mathbf{P} - \mathbf{N}_p\right) \times \mathbf{W} \times (-\mathbf{1})^{\mathbf{Neg}} \tag{1}$$

where P represents the number of active words, N_p refers to the number of passive words, and Neg denotes the count of negation words (such as "not," "no," and other similar terms). The sentiment score for each sentence is adjusted by multiplying it by $(-1)^{Neg}$, reflecting the principle that double negation leads to affirmation [48]. W is the weights corresponding to adverbs of degree. If there are no adverbs of degree, the weight is considered as 1. Generally, a sub-sentence contains only one adverb of degree. If there are multiple adverbs of degree, the weight corresponding to the first matched adverb is used.

$$\mathbf{Review'sscore} = \sum \mathbf{Sentence'sscore} \tag{2}$$

$$\mathbf{CS} = \frac{\mathbf{Review'sscore} - \mathbf{Review'sscore_{min}}}{\mathbf{Review'sscore_{max}} - \mathbf{Review'sscore_{min}}} \times \mathbf{100} \tag{3}$$

where CS is represented as a numerical value quantified on a scale from 0 to 100, derived from text sentiment analysis.

3.2 BCN construction

3.2.1 Node determination

In this study, co-movement describes the relationship between CS and SP, while the co-movement symbol represents both the direction and magnitude of their interaction. The coarse graining method in symbol dynamics and time series involves defining the co-movement sequence as Lp_i, which represents the product of ΔCS (the first-order difference of CS) and ΔSP (the first-order difference of SP) [49]. Avg+ signifies the average of $\Delta CS \times \Delta SP$ when their product is positive, while Avg− donates the average of the product is negative. Subsequently, Lp_i is subdivided into the following five cases, as shown in Equation 4.

$$\mathbf{Lp_i} = \begin{cases} \mathbf{Y}, \Delta\mathbf{CS} \times \Delta\mathbf{SP} \geq \mathbf{Avg+} & \mathbf{Violent\ co-movement\ in\ the\ same\ direction} \\ \boldsymbol{y}, \mathbf{0} < \Delta\mathbf{CS} \times \Delta\mathbf{SP} < \mathbf{Avg+} & \mathbf{Slight\ co-movement\ in\ the\ same\ direction} \\ \boldsymbol{o}, \Delta\mathbf{CS} \times \Delta\mathbf{SP} = \mathbf{0} & \mathbf{No\ co-movement} \\ \boldsymbol{n}, \mathbf{Avg-} < \Delta\mathbf{CS} \times \Delta\mathbf{SP} < \mathbf{0} & \mathbf{Slight\ co-movement\ in\ opposite\ directions} \\ \boldsymbol{N}, \Delta\mathbf{CS} \times \Delta\mathbf{SP} \leq \mathbf{Avg-} & \mathbf{Violent\ co-movement\ in\ opposite\ directions} \end{cases} \tag{4}$$

Thus, the co-movement between CS and SP can be represented by a continuous symbol sequence, and the corresponding symbol sequence is given in Equation 5.

$$\mathbf{Lp_i} = \left(\mathbf{Lp_1}, \mathbf{Lp_2}, \ldots, \mathbf{Lp_w}\right)\left(\mathbf{Lp_i} \in (\mathbf{Y}, \mathbf{y}, \mathbf{o}, \mathbf{n}, \mathbf{N}), \mathbf{w} = \mathbf{4}\right) \tag{5}$$

Given that stock trading typically spans five consecutive days, with statutory holidays excluded from consideration, we define a node as the co-movement of four consecutive days. w = 4 ensures the balance of the time window, though its value can be adjusted based on the sample size in this study. The symbol sequence Lp_i serves as the data sliding window with a step size of 1, and the coarse-grained symbol sequence is converted into partially overlapping symbol intervals, each representing different co-movement modes.

3.2.2 Determination of edges and weights

Using the coarsened data, edges in the co-movement network are established by connecting the previous co-movement mode combination to the neighboring subsequent co-movement mode combination. The number of transitions between two combinations of the same mode function as the edge weight. Therefore, a weighted directed BCN is constructed to depict the co-movement modes between CS and SP.

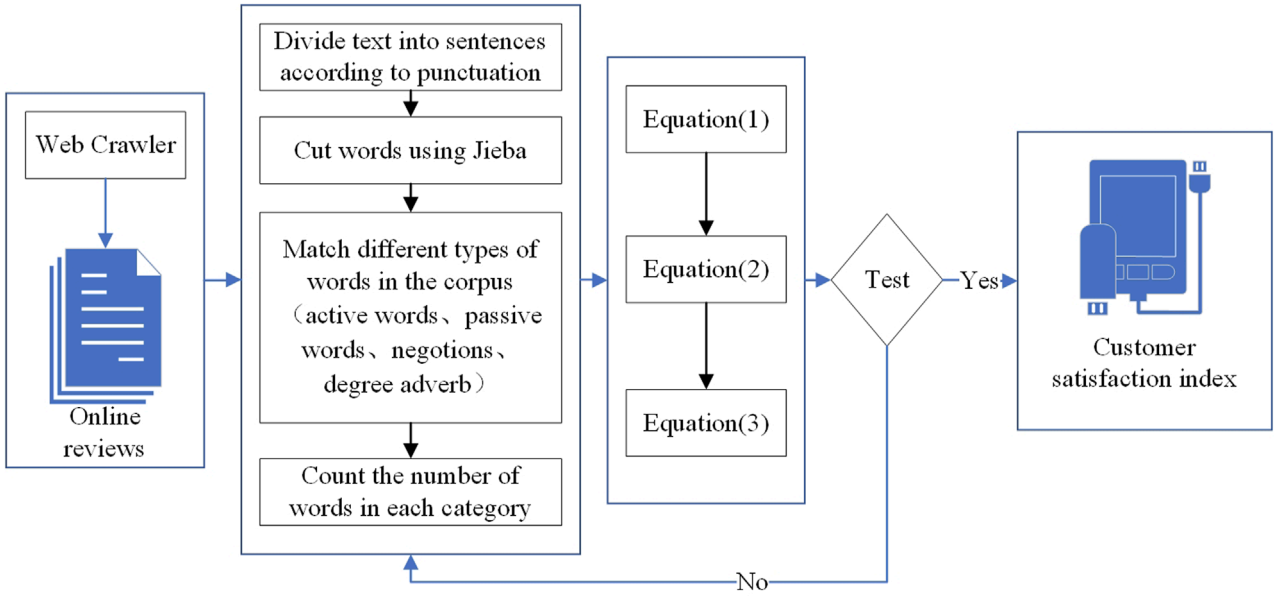


FIGURE 1
Text sentiment analysis flowchart.

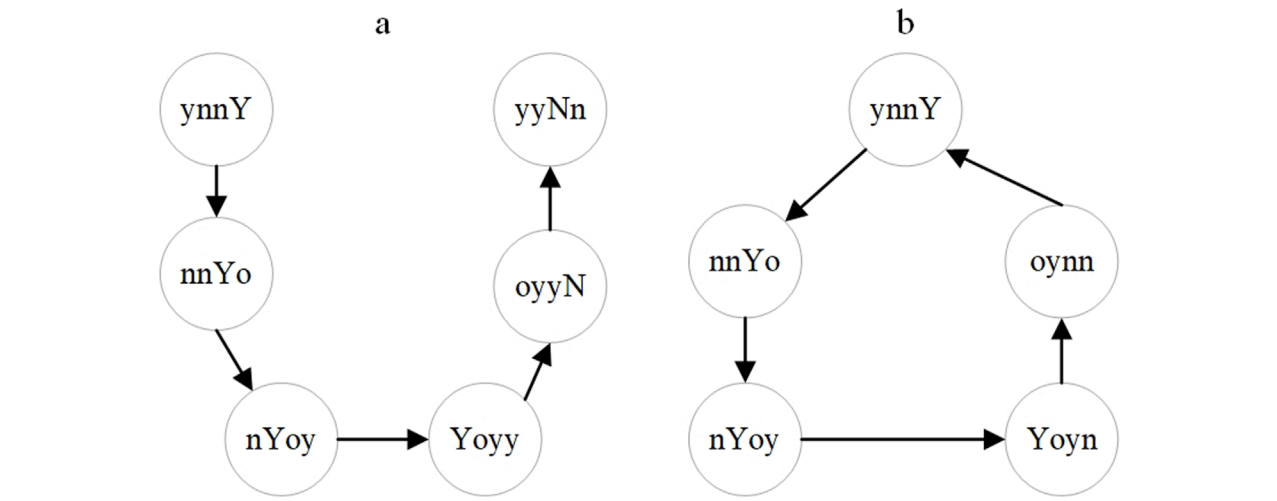


FIGURE 2
Network structure corresponding to the latest data array.

If the symbol sequence mapped for the time series is {ynnYoyyN n}, the corresponding latest node is {yyN n}. Figure 2A illustrates the network structure, where the number of nodes equals the count of co-movement modes. However, if the symbol sequence mapped for the time series is {ynnYoynnY}, and the latest node is {ynnY}, the number of nodes is less than the count of co-movement modes, as depicted in Figure 2B. In practical network construction, when the sample size is sufficient, it often resembles the structure shown in Figure 2B. If the sample size is limited, one can reduce the use of symbols or symbol combinations (i.e., reduce w) to construct the BCN.

3.3 Feature extraction

We derive BCN's features across four key dimensions to guide the subsequent predictions, with the specific features and indicators detailed in Table 1.

3.4 Co-movement prediction model

Co-movement forecasting is the prediction of the next day's CS and SP co-movement symbol. For instance, if we observe a co-movement mode on day t, denoted as {*ynny*}, forecasting the

TABLE 1 Outlines the framework of the feature extraction model for BCN.

Dimension	Indicators	Formulas
Influence	Node strength	$S_{i_out} = \sum_{j \in N_i} W_{ij}$, $S_{i_in} = \sum_{j \in N_i} W_{ji}$, $S_i = S_{i_out} + S_{i_in}$ [49] where S_i is node strength of node i, N_i is the set of neighbors of node i, and W_{ij} is the weight of node i to node j
	Strength distribution	$P(s) = \frac{S_i}{\sum_{i=1}^{N} S_i}$ where N is the total number of nodes, $P(s)$ is strength distribution
Clustering	Clustering coefficient	$C_i = \frac{E_i}{k_i(k_i-1)}$ [49] where C_i is the clustering coefficient of node i, k_i is the degree of node i, and E_i is the number of actual edges between the neighboring nodes of node i
	Weighted Clustering coefficient	$C_{i_weighted} = \frac{1}{S_i(k_i-1)} \sum_{(j,k)} \frac{w_{ij}+w_{ik}}{2} a_{ij}a_{ik}a_{jk}$ where a_{ij} indicates whether there is a link between node i and node j: 1 if a link exists, and 0 otherwise
Periodicity	Betweenness centrality	$W_k = \sum_{(i,j)} \frac{C_k(i,j)}{C(i,j)}$ [15] where W_k is the betweenness centrality of node k. $C(i,j)$ is the total number of all shortest paths between node i and node j. $C_k(i,j)$ is the number of paths through node k
	Average path length	$L_{avg} = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$ [15] where d_{ij} is the number of edges in the shortest path between nodes i and j
Orderliness	Standards Network Strength Entropy	$G = \frac{-\sum_{i=1}^{N} P(s) ln P(s)}{lnN}$, $G_{i_in} = \frac{-P(s)_{ji} ln P(s)_{ji}}{ln n_{i_in}}$, $G_{i_out} = \frac{-P(s)_{ij} ln P(s)_{ij}}{ln n_{i_out}}$ [58] where G presents the standards network strength entropy. G_{i_in} and G_{i_out} are the standards strength in-entropy and out-entropy of node i. n_{i_in} is the number of nodes entering node i. n_{i_out} is the number of nodes that node i points to. $P(s)_{ji} = \frac{w_{ji}}{\sum_{j=1}^{n} w_{ji_in}}$, $P(s)_{ij} = \frac{w_{ij}}{\sum_{j=1}^{N} S_{ij}}$

TABLE 2 Node strength and strength distribution of each co-movement mode.

No.	Node	S	Ratio (%)	No.	Node	S	Ratio (%)
1	nnnn	98	4.34	11	nyyy	68	3.01
2	nnny	80	3.54	12	ynny	67	2.96
3	ynyn	78	3.45	13	nnyy	66	2.92
4	ynnn	78	3.45	14	nnyn	66	2.92
5	nyny	78	3.45	15	yyyn	62	2.74
6	yyyy	76	3.36	16	nynn	62	2.74
7	ynyy	74	3.27	17	yyYy	14	0.62
8	yyny	70	3.10	18	yYyn	14	0.62
9	yynn	70	3.10	…	…	…	…
10	nyyn	70	3.10	268	YYny	1	0.04

mode for day $t+1$ relates to predicting the mode {*nny*?}. Prior studies have indicated that the total occurrence time of different co-movement modes follows an approximately linear distribution [49]. This implies that with sufficient data, new co-movement modes will not arise in the short term. Hence, we utilize mode transformation relationships for forecasting.

3.4.1 BCN optimization based on entropy principal

To ensure the accuracy of predictions, it is essential to adjust the network structure to enhance its clarity, particularly in cases where chaotic nonlinear systems exhibit strong randomness. Therefore, the partial edge weights are adjusted to improve its predictability based on entropy optimization. Entropy optimization is a method for finding the optimal solution under given constraints by maximizing or minimizing an entropy-based objective function, leveraging the mathematical properties of entropy and principles of information theory [50]. Entropy, as an optimization and correction tool, is used to optimize the return statistics that do not follow Gaussian distribution, and to correct the geometric Brownian motion model by relaxing the assumption of lognormal distribution [51, 52]. Furthermore, entropy optimization is utilized in identifying topological changes in complex networks, assessing network consistency, and reducing knowledge distillation loss in deep neural networks by adjusting network entropy [53–55]. The use of entropy is prevalent in economics, with applications ranging from risk measurement and portfolio optimization to the design of combined economic emission dispatch optimization algorithms, all of which contribute to enhancing the scientific basis for decision-making [51, 56, 57]. The aforementioned studies demonstrate the broad applicability and notable effectiveness of entropy optimization across various economic domains.

The standard network strength entropy is employed as a measurement tool, which extends Shannon entropy to quantify the uncertainty in the distribution of node or edge strength [58]. This measure is standardized to enhance its applicability to network analysis. Building upon this, the approach constructs a bi-objective entropy optimization model that aims to minimize network distortion and simultaneously adjusts the weights of selected edges to reduce the network's output entropy (see G_1), thereby achieving overall network optimization. Minimizing network distortion ensures that the adjustment of edge weights has minimal impact on the network's topological structure, which is mathematically represented by G_2. The selected edges set refers to the edges connected to nodes with higher recognizability, which

are typically associated with lower entropy values. These edges are crucial for reducing network out-entropy while preserving the network's essential features. Based on these objectives, the following model is proposed (see Equation 6):

$$G_1 = \min_x \left\{ \frac{\sum_i^N G_{i_{out}}}{N} \right\}$$
$$G_2 = \min_x \left\{ \frac{\left|\sum_{i=1}^N \Delta S_i\right|}{N} + \frac{\left|\sum_{i=1}^N \Delta C_{i_{weighted}}\right|}{N} \right\} \quad (6)$$

Where x represents the adjustment applied to the weight of the edge.

3.4.2 RC-BCN model

In the process of constructing a BCN, it is crucial to consider two scenarios for the latest nodes:

(1) If the latest node $V_i^{(t)}$ has no out-neighbor node in BCN (see Figure 2A), then the future node that appears at $t+1$ can be estimated according to the trend extrapolation method [59]. That is, if $V_i^{(t)} = \{yyny\}$, then $V_{i\to j}^{(t+1)} = \{ynyy\}$. In fact, this situation can be avoided when building the BCN.

(2) If $V_i^{(t)}$ has out-neighbor nodes in BCN (see Figure 2B), the set of out-neighbor nodes is defined by Equation 7.

$$V_{i\to j}^{(t+1)} = \left\{ V_{i\to j_1}^{(t+1)}, V_{i\to j_2}^{(t+1)}, \ldots, V_{i\to j_n}^{(t+1)} \right\}, j_1, j_2 \ldots, j_n \in [1, N] \quad (7)$$

Based on the preceding analysis (3.2.2), the future nodes at $t+1$ are one of the elements in $V_{i\to j}^{(t+1)}$. There are three common rules to determine the exact nodes.

Rule 1: The principle of the greatest Node Similarity. Link prediction techniques are applied, where the probability of a potential connectivity between nodes is determined by computing the similarity between two nodes. The node with the highest similarity is selected as the predicted node, and in case of tie, preference is given to the relatively most recent node.

Given a pair of nodes (v_i, v_j), the rules for calculating the node similarity are as follows: If the BCN exhibits significant clustering, a local feature-based similarity formula is applied (see Equation 8); otherwise, a global-based similarity formula is utilized (see Equation 9) [42].

$$S_{ij}^{JC} = \frac{|\Gamma(i) \cap \Gamma(j)|}{|\Gamma(i) \cup \Gamma(j)|} \quad (8)$$

where $\Gamma(i)$ and $\Gamma(j)$ are the sets of out-neighbor nodes of v_i and v_j respectively.

$$S_{ij}^{WJC} = \frac{\sum_{x \subset \Gamma(i) \cap \Gamma(j)} \left(w_{ix} + w_{xj} \right)}{\sum_{x \subset \Gamma(i) \cup \Gamma(j)} \left(w_{iy} + w_{yj} \right)} \quad (9)$$

where w_{ix}, w_{xj}, w_{iy} and w_{yj} are edge weights, respectively. This formula takes into account the weight of the edge on the basis of Equation 8.

If the maximum value of either S_{ij}^{JC} or S_{ij}^{WJC} is 0, it is not possible to predict future nodes. In this case, Rule 2 or Rule 3 are employed for combined prediction.

Rule 2: The principle of the greatest node strength, that is, select the node with the largest node strength from $V_{i\to j}^{(t+1)}$. In case of tie, the relatively most recent node is selected.

Rule 3: The principle of the greatest connection strength, that is, select the node with the largest connection strength from $V_{i\to j}^{(t+1)}$. In the event of a draw, the relatively most recent node is selected.

Following these rules, co-movement modes prediction function was established, as shown in Equation 10.

$$\hat{V}_{i\to j}^{(t+1)} = f\left(V_i^{(t)} \right) + \xi_t, t \in [0, l] \quad (10)$$

where $f(x)$ is the composite function comprising six co-movement prediction models: M-1 (Trend extrapolation method and Rule 1-Equation 8 and Rule 2), M-2 (Trend extrapolation method and Rule 1-Equation 8 and Rule 3), M-3 (Trend extrapolation method and Rule 1-Equation 9 and Rule 2), M-4 (Trend extrapolation method and Rule 1-Equation 9 and Rule 3), M-5 (Trend extrapolation method and Rule 2), and M-6 (Trend extrapolation method and Rule 3). ξ_t is the error.

3.4.3 Evaluation of model prediction accuracy

The co-movement prediction of CS and SP refers to forecasting both the direction and strength of their co-movement. Hence, model prediction accuracy is assessed in terms of two aspects: accurately predicting both the co-movement direction and strength, or solely the co-movement direction. The evaluation indicators for accurately predicting both co-movement direction and strength are defined in Equations 11, 12.

$$D_1 = \frac{1}{L} \sum_{t=1}^{L} a_t \quad (11)$$

$$a_t = \begin{cases} 1, & V_{i\to j}^{(t+1)} = \hat{V}_{i\to j}^{(t+1)} \\ 0, & V_{i\to j}^{(t+1)} \neq \hat{V}_{i\to j}^{(t+1)} \end{cases} \quad (12)$$

The evaluation indicators for accurately predicting the co-movement direction are given in Equations 13, 14.

$$D_2 = \frac{1}{L} \sum_{t=1}^{L} b_t \quad (13)$$

$$b_t = \begin{cases} 1, & V_{i\to j}^{(t+1)} \cong \hat{V}_{i\to j}^{(t+1)} \\ 0, & V_{i\to j}^{(t+1)} \neq \hat{V}_{i\to j}^{(t+1)} \end{cases} \quad (14)$$

$V_{i\to j}^{(t+1)} \cong \hat{V}_{i\to j}^{(t+1)}$ signifies the predicted and actual symbols have matching directions.

4 Experiments

4.1 Data preparation

A set of experiments was conducted using JD.com, a prominent B2C online shopping platform. According to a 2022 Tencent report, electronic products on JD.com accounted for over 60%. Hence, 44 high-selling electronic products were selected from JD. com's main

product categories, encompassing five renowned brands: Huawei, Apple, Samsung, Xiaomi, and OPPO.

A total of 92,722 CORs were collected, spanning from 9 January 2018, to 15 August 2022. The collected CORs include the consumers' review and the date. Textual sentiment analysis resulted in a 1,645-day CS dataset. Given that multiple data points exist for the same day, we use the daily average to more accurately reflect the overall CS for that day. The closing price from 9 January 2018, to 15 August 2022, was obtained from the Straight Flush website.

After data merging, 1,135 sets of data were obtained. Subsequently, a Spearman correlation coefficient test was conducted, revealing a significant positive correlation between them. The summary statistics indicate that the data do not follow a normal distribution (see Supplementary Appendix A). This suggests that the relationship between the variables is complex and warrants further analysis using nonlinear methods. To further process the data, coarse-graining and sliding window techniques were employed, resulting in the generation of 268 nodes and 1,134 edges.

4.2 CS and SP co-movement network construction

Figure 3 displays the CS and SP Co-movement Network (CS-SPCN). It shows obvious differences in the influence of the co-movement modes, and the clustering between different modes.

Furthermore, as the data volume expands, the increase in node count exhibits a diminishing trend (see Figure 4). The temporal intervals between the occurrence of new co-movement modes and those of old modes eventually grow longer (see Figure 5). This implies that future modes are likely to be once-emerged modes. It also demonstrates that JD. com's CS-SPCN adheres to the application premise of the co-movement mode prediction model developed in this study.

4.3 Feature extraction

4.3.1 Influence

(1) The co-movement between CS and SP is primarily slightly negative. Specifically, there were 98 instances of opposite-direction co-movement over four consecutive days, while only 76 instances of same-direction co-movement occurred (see Table 2). This reflects the risk of reverse reactions and potential misalignment between CS and SP.

(2) The CS-SPCN exhibits an uneven distribution, characterized by significant disparities in nodes and edges. The initial 16 nodes demonstrate substantial node strength, accounting for a cumulative strength distribution of 51.46%. In contrast, there are 113 nodes with a node strength of two or less, representing only 9.96% of the total strength (see Figure 6A). Figure 6B demonstrates that the cumulative strength distribution approximately follows a power-law distribution, indicating that the CS-SPCN is a scale-free network. There are primarily 16 manifestations of co-movement modes. These 16 co-movement modes frequently transition to other modes, or vice versa, demonstrating a high frequency of shifts between modes. This suggests that market reactions are nonlinear and complex, and companies must focus on the most influential factors, concentrating their resources on understanding core sentiment and market responses.

4.3.2 Clustering

The CS-SPCN exhibits a certain degree of clustering, with some nodes having strong connections, forming small and tightly-knit communities or subgroups. There are 29 nodes with a non-zero clustering coefficient. Both {*yyyy*} and {*nnnn*} have greater clustering coefficients and strength (see Figure 7). The highest clustering coefficient value observed is 0.5, suggesting a generally modest level of network cluster. Additionally, the weighted clustering coefficients of the nodes {*nnnn*}, {*Ynnn*} and {*ynnn*} exceed 0. This suggests that co-movements are interrelated, and companies should consider the interactions of multiple co-movements, rather than focusing on a single direction, in decision-making.

4.3.3 Periodicity

(1) Nodes with violent co-movement symbols exert substantial network influence. In terms of transition capabilities, {*ynny*, *nnyY*, *Nynn*, *YyYN*, *YYyY*} are the main transformation hubs. Businesses and investors should proactively identify market signals and focus on key moments to leverage critical decision-making opportunities. Figure 8A indicates a relatively even distribution of node betweenness centrality. Figure 8B shows that the top 26 nodes have a 25% share of cumulative betweenness centrality, indicating that 9.7% of the nodes perform 25% of the transit function. Some nodes with low strength play essential hub roles, as evidenced by the circled points in Figure 8C.

(2) The co-movement between CS and SP exhibits a cyclical pattern, with a periodicity of 8 days. The CS-SPCN demonstrates an average path length of 7.696. Given a mode generation sliding unit of 1 day, transitions between modes span approximately 8 days, thus accentuating the periodic feature of the co-movement. Companies should develop forecasting models to adjust decisions promptly in response to market cycles.

4.3.4 Orderliness

Overall, the network exhibits a high degree of randomness and disorder, as evidenced by the standard network strength of the BCN exceeding 0.8, which suggests that predicting the next node in the network without relying on local information is highly challenging. This also indicates that businesses operate in a highly uncertain environment, where both businesses and investors face the risk of being unable to accurately predict future market changes, thereby impacting the effectiveness and precision of their decisions. This study addresses the challenge of predicting the nodes at $t+1$ based on the nodes at t. Networks with lower out-entropy demonstrate higher predictability, as reduced out-entropy reflects less uncertainty in the outgoing connections of nodes. Analyzing the local characteristics reveals that the average in-entropy and out-entropy of each node follow specific patterns. Particularly, nodes set with violent co-movement at the end state, exhibit the lowest average in-entropy and out-entropy (see Table 3). As a result,

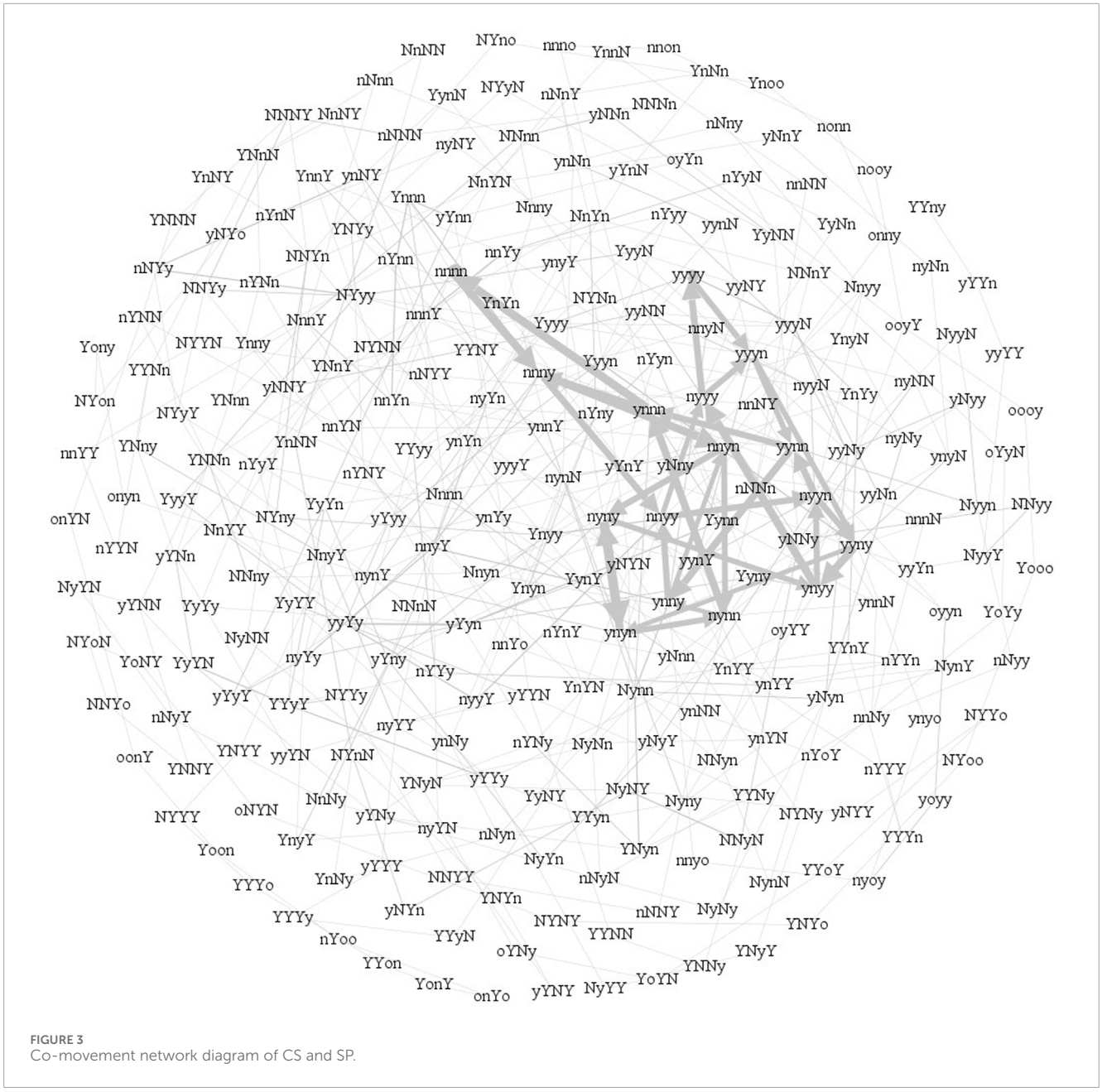


FIGURE 3
Co-movement network diagram of CS and SP.

adjusting the edge weights of nodes exhibiting violent co-movement at the end state proves effective in reducing the average out-entropy of the BCN.

To enhance the network's predictability, the edge weight of the node set with violent co-movement at the terminal state is adjusted to x. To minimize network distortion and effectively reduce network out-entropy, the value of x is constrained within the range [2,7.696]. To accurately obtain the Pareto frontier in a bi-objective optimization problem, the Eel and Grouper Optimizer (EGO) algorithm is employed. EGO is a novel meta-heuristic algorithm inspired by the symbiotic interactions and foraging strategies of eels and groupers in Marine ecosystems [60]. The algorithm's flexible transition between the exploration and development phases, along with its population-based nature, effectively reduces the risk of converging to local optima. The Pareto frontier is calculated using EGO, as illustrated in Figure 9. Excluding cases where $G_2 > 6$ leads to significant network distortion, the analysis categorizes the network into three groups based on decision-makers' preferences regarding the level of network distortion.

Given that $4 < G_2 < 6$, the decision-maker shows a preference for network distortion. For $2 < G_2 < 4$, distortion neutrality is observed. If $0 < G_2 < 2$, the decision-maker demonstrates aversion to distortion. Three optimal solutions are selected from these three acceptable regions for subsequent prediction.

4.4 Experimental analysis

Sample of model estimation. The growth of a business usually follows an "S" curve. In addition, the network structure exhibits

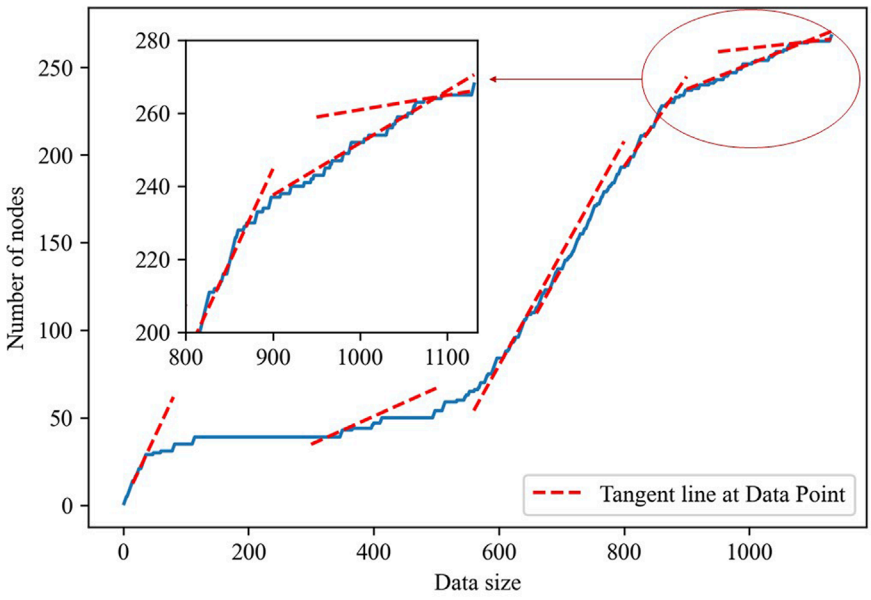


FIGURE 4
Relationship between Data Size and Number of Nodes.

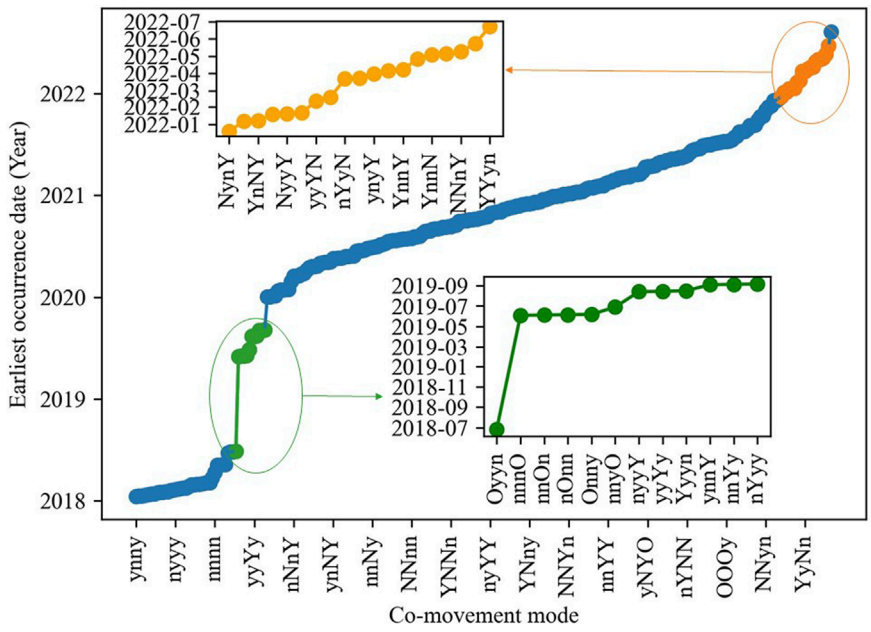


FIGURE 5
Relationship between Co-movement Modes and Earliest Occurrence Times.

annual instability in the post-pandemic period. By utilizing up-to-date data, more precise predictions can be generated. Hence, a sample of co-movement modes from 2022 is employed to evaluate the accuracy of the model.

Implementation details. To assess the effectiveness of BCN optimization, we conducted four experimental sets with varying distortion preferences and baseline model, comparing the prediction accuracy of each model. Experiment A0 serves as the baseline model, where the BCN is not optimized. Experiment A1 represents the predictive model under distortion aversion ($x = 2.21$), Experiment A2 corresponds to distortion neutrality ($x = 4.36$), and Experiment A3 reflects a distortion preference ($x = 6.31$). These experiments were designed to investigate the impact of varying distortion attitudes on predictive accuracy, thereby validating the effectiveness of BCN optimization. Furthermore, given the limited frequency distribution of violent co-movements, the analysis focuses solely on

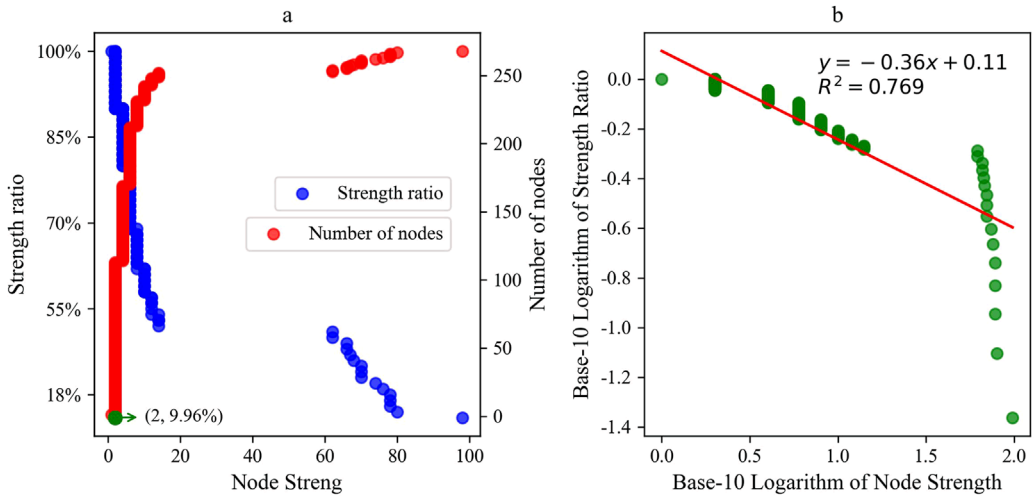


FIGURE 6
Relationship of Node Strength with Strength Ratio and Number of Nodes. Strength Ratio represents the ratio of the cumulative strength of the nodes to the sum of all nodes' strength.

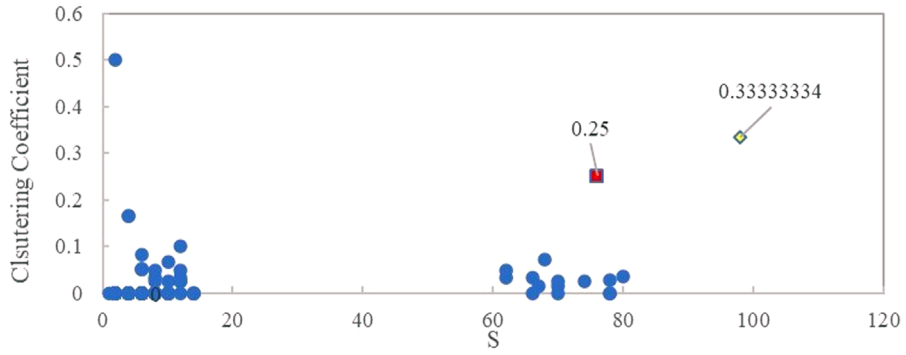


FIGURE 7
Relationship between Node Strength and Clustering Coefficient.

forecasting the direction of co-movement rather than its degree. The accuracy of the prediction model is evaluated using D_2. We randomly selected 200 samples from the dataset and calculated the accuracy of the prediction results.

Figure 10A0 shows that M-5 and M-6 demonstrate higher prediction accuracy compared to other models, suggesting that node strength and edge weight are the primary factors influencing the co-movement between CS and SP, while the influence of node similarity appears to be negligible. By comparing with Figure 10A0, it is evident that the prediction accuracy of M-5 and M-6 is significantly improved by entropy optimization, whereas the impact on the prediction accuracy of M-1 to M-4 is less pronounced. This is primarily because the optimization of the network has a direct and significant impact on the strength of nodes and edges, but does not significantly affect the similarity between nodes. Furthermore, as the forecasting length (FL) increases, the prediction accuracy tends to converge around 50%–60%. The prediction accuracy of M-6 is nearly identical under the conditions of distortion aversion and neutrality, with both exceeding 80% (see Figures 10A1, A2). The prediction accuracy of M-5 and M-6 can approach 90% under the distortion preference condition (see Figure 10A3).

Compared to RC-BCN alone, the prediction accuracy of our proposed RC-BCN based on entropy optimization is, on average, improved by 15.5% under M-6. To further elucidate the contribution of this study, we compare the method proposed herein with the prediction results obtained from the BP neural network and LSTM models. For the problems addressed in this paper, both the BP neural network and LSTM models use the autoregressive method to predict the values of CS and SP, respectively. The co-movement direction of CS and SP is then calculated based on the prediction results, with the methods denoted as BP_1 and $LSTM_1$. Furthermore, the direction of co-movement can be directly classified and predicted, denoted as BP_2 and $LSTM_2$. The parameter settings for the four models are presented in Table 4 and Figure 11 illustrates that M-5 and M-6 outperform both BP and LSTM in terms of prediction accuracy.

We employ M-5 and M-6 to predict the next 4 days co-movement modes under the different preferences of decision makers. The latest co-movement mode is $\{YYny\}$, originating from

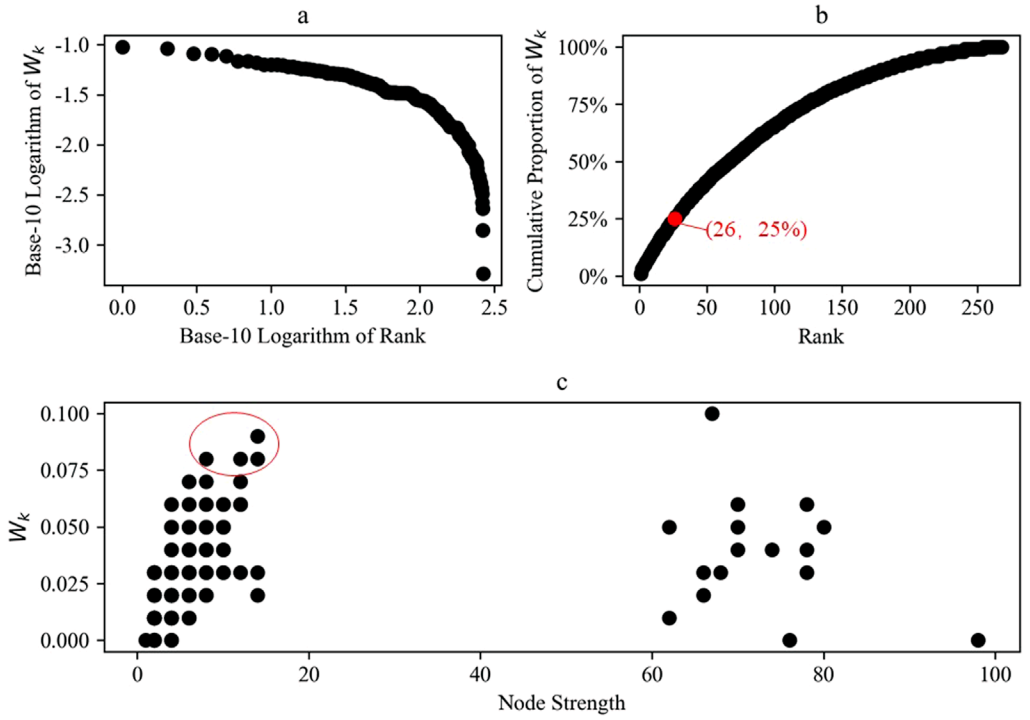


FIGURE 8
Betweenness centrality distribution. Rank refers to the nodes ordered according to W_k. Cumulative proportion of W_k represents the ratio of the cumulative W_k of the nodes to the sum of all nodes' W_k.

TABLE 3 Network entropy.

Items	A	B	A + B
Total in-entropy	55.46	71.97	127.43
Average in-entropy	0.45	0.50	0.48
Percentage of Nodes with In-Entropy = 0	53.23%	45.14%	48.88%
Total out-entropy	54.46	71.30	125.77
Average out-entropy	0.44	0.50	0.47
Percentage of Nodes with Out-Entropy = 0	54.84%	45.14%	49.63%
Number of nodes	124	144	268

[a]A node set exhibiting violent co-movement at its terminal state; B A node set exhibiting slight co-movement at its terminal state.

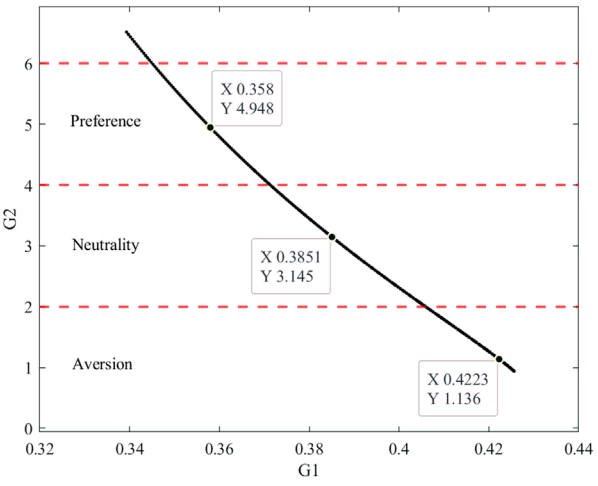


FIGURE 9
Pareto Frontier in the bi-objective optimization.

15 August 2022. The co-movement modes for the last 4 days are $\{Ynyy, nyyy, yyyy, yyyn\}$ across different scenarios.

5 Discussion

The co-movement forecast between CP and SP could assist publicly listed e-commerce companies in financing decisions and investors in making investment decisions. Motivated by it, we explore the forecast approach of the co-movement modes between the two based on complex network and entropy optimization. Specially we project the time series data into the phase space, and enhance the prediction through rule combination and entropy optimization instructed by features extracted from network topology indicators. We conducted four experimental sets, with results indicating the model's successful predictive performance.

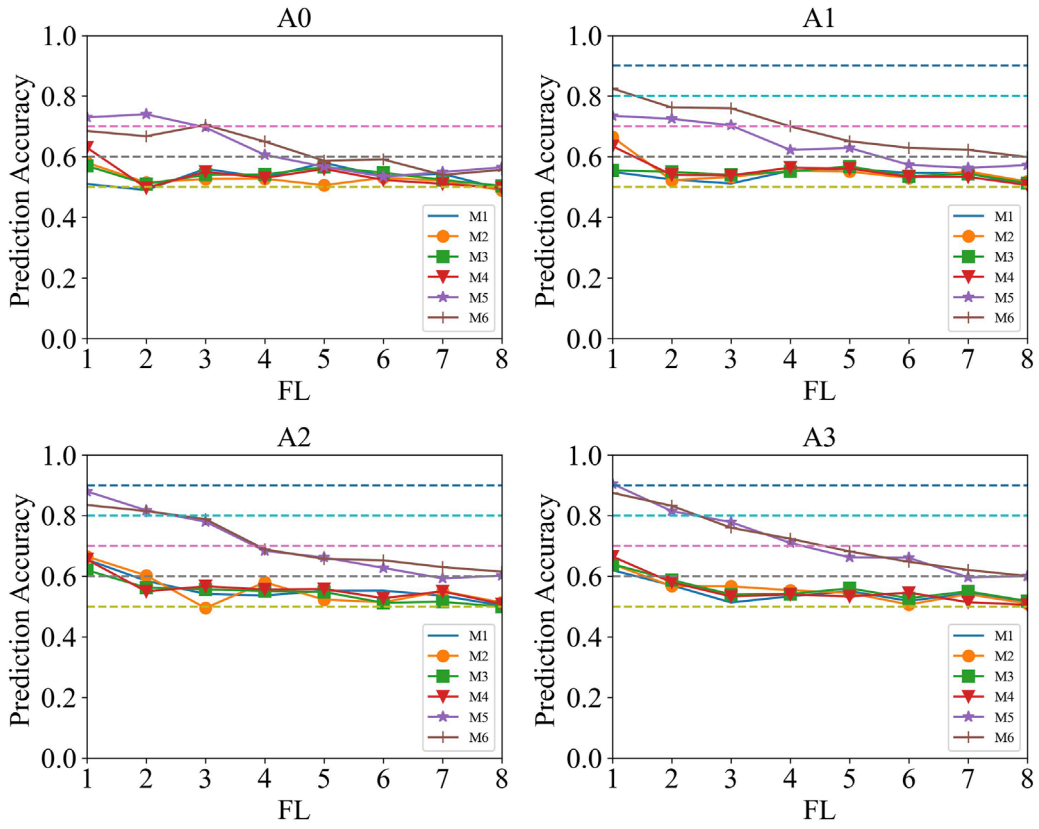


FIGURE 10
Model prediction accuracy under different FL.

TABLE 4 Model parameter setting.

	Input layer dimension	Output layer dimension	Number of neurons in hidden layer	Batch size	Training rounds	Learning rate	Validate frequency
BP_1	1	1	50	80	200	0.001	5
$LSTM_1$							30
BP_2	2						5
$LSTM_2$							30

This paper provides a new perspective for the prediction of chaotic nonlinear systems and expands the application of entropy optimization within the field of e-commerce.

5.1 Result analysis

This study's proposed approach for co-movement prediction in CS and SP yielded favorable results. Particularly the study innovatively puts forward entropy optimization guided by network features to enhance co-movement prediction. In our experiment, we construct CS-SP CN using time series data, and combine the designed forecasting rules with network features to form bivariate co-movement forecasting based on rule combination for the first time. We call this novel approach RC-BCN. Comprehensive experiments show that the probability of establishing connections between highly similar modes is minimal in CS-SPCN. This may be attributed to the influence of nonlinear relationships, external variables, or temporal factors within CS-SPCN. The connection among co-movement modes is primarily influenced by node strength and edge weight in CS-SPCN. M-5 involves identifying the node with the strongest strength among the external nodes corresponding to the target node as the prediction outcome. In the case of M-6, the prediction outcome is determined by selecting the mode with the highest connection weight among the externally connected modes of the co-movement mode. These two

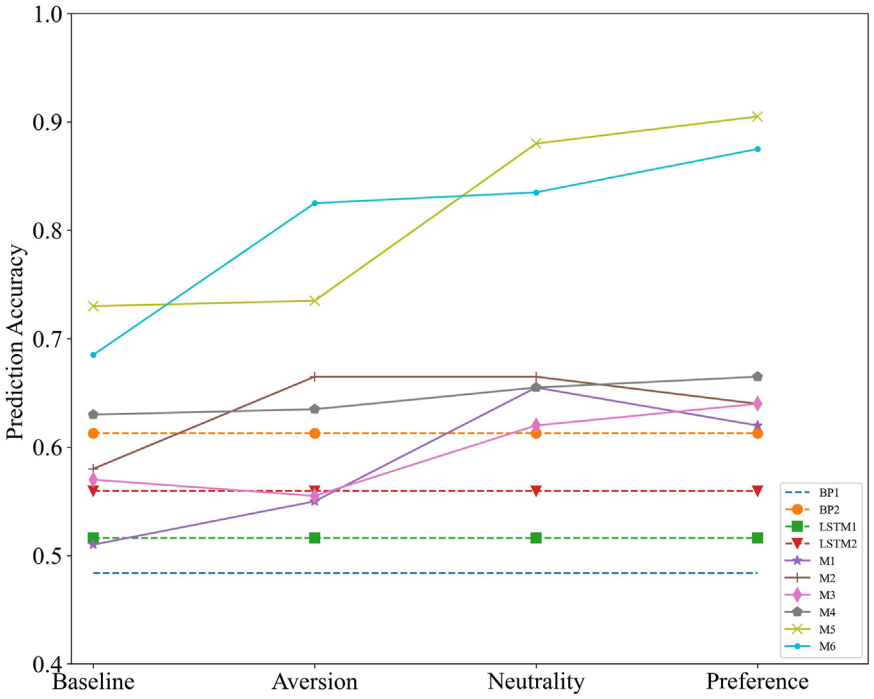


FIGURE 11
Prediction accuracy across different models when FL = 1.

models perform well. This demonstrates the effective integration of complex networks and entropy optimization into time series analysis, resulting in enhanced performance.

The sample size and data source may limit the generalizability of the results. Nevertheless, the JD platform and the electronic product categories selected in this study are representative, providing valuable insights into typical consumer behavior and financial market dynamics. Moreover, the robustness of the results has been verified using randomized experiments and sensitivity analysis (detailed in Supplementary Appendix B). To extend these findings, future research could explore enterprises across various industries, beyond just e-commerce, further validating and enriching the applicability of the proposed approach.

5.2 Theoretical and practical implications analysis

From a theoretical perspective, three outcomes are worth mentioning. Firstly, we delve into the co-movement between CS and SP on a finer time scale, providing valuable insights to support investor decision-making and corporate financing strategies. Secondly, we introduce a novel co-movement prediction model based on BCN and rule combination. Thirdly, we extend the application of entropy optimization in forecasting the co-movement between CS and SP, offers a new perspective on their relationship, and contributes to the development of entropy theory in economics.

Regarding practical contributions, the co-movement forecasting provides valuable insights for corporate financing and investor decision-making by revealing the intricate relationship between CS and SP. For companies, leveraging this market feedback can help optimize financing strategies and assess potential risks. If negative sentiment is predicted to lead to a decline in stock prices, enterprises may opt to raise funds during periods of more favorable market sentiment, thereby reducing financing costs. For investors, understanding the predicted co-movement patterns, especially sustained opposite trends, can help them make more informed investment decisions and adopt appropriate risk mitigation strategies, ensuring cautious investment in volatile market conditions.

6 Conclusion and future work

In this study, we integrated complex network theory and entropy optimization into time series analysis, introducing a novel RC-BCN method to forecast the co-movement between CS and SP, which performs well. Additionally, the co-movement between CS and SP of JD is primarily driven by node strength and edge weight, suggesting that nodes and node pairs with higher frequencies of past appearances are more likely to reappear in the future. This indicates that the co-movement between CS and SP is influenced not only by the continuity of existing market behaviors and interaction structures, but also reflects the Matthew effect, that is, frequently occurring co-movement patterns are more likely to continue to dominate the co-movement in the future. Moreover, this study contributes a novel perspective and methodological framework for examining bivariate co-movement, offering significant implications for enterprises and investors in optimizing financing and investment strategies.

Similar to many studies, this research is not without its limitations. First, this study examines co-movement between variables in the same or opposite directions, but does not distinguish whether the co-movement in the same direction is positive or negative. Second, while many variables affect corporate financing and investment, this research focuses solely on two-variable co-movement prediction. Future studies will extend this method to forecast the co-movement among additional economic variables to enhance the generalizability of the findings. By incorporating the specific co-movement directions of each variable, both businesses and researchers can gain deeper insights into the interdependent relationships among complex economic systems in e-commerce.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

MW: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Writing–original draft, Writing–review and editing. RK: Conceptualization, Supervision, Writing–review and editing. JL: Methodology, Writing–review and editing. WH: Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors appreciate the valuable comments of the editors and reviewers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1557361/full#supplementary-material

References

1. Shapiro AH, Sudhof M, Wilson DJ. Measuring news sentiment. *J Econom* (2022) 228(2):221–43. doi:10.1016/j.jeconom.2020.07.053

2. Wang H, Shen H, Li S. ESG performance and stock price fragility. *Finance Res Lett* (2023) 56:104101. doi:10.1016/j.frl.2023.104101

3. Chaudhari K, Thakkar A. Data fusion with factored quantization for stock trend prediction using neural networks. *Inf Process Manag* (2023) 60(3):103293. doi:10.1016/j.ipm.2023.103293

4. Eachempati P, Srivastava PR, Kumar A, Munoz de Prat J, Delen D. Can customer sentiment impact firm value? An integrated text mining approach. *Technol Forecast Soc Change* (2022) 174:121265. doi:10.1016/j.techfore.2021.121265

5. Abdelfattah BA, Darwish SM, Elkaffas SM. Enhancing the prediction of stock market movement using neutrosophic-logic-based sentiment analysis. *J Theor Appl Electron Commer Res* (2024) 19(1):116–34. doi:10.3390/jtaer19010007

6. Nguyen H, Calantone R, Krishnan R. Influence of social media emotional word of mouth on institutional investors' decisions and firm value. *Manage Sci* (2020) 66:887–910. doi:10.1287/mnsc.2018.3226

7. Li J, Zhang Y, Li J, Du J. The role of sentiment tendency in affecting review helpfulness for durable products: nonlinearity and complementarity. *Inf Syst Front* (2023) 25(4):1459–77. doi:10.1007/s10796-022-10292-3

8. Bennett B, Stulz R, Wang Z. Does the stock market make firms more productive? *J Financ Econ* (2020) 136(2):281–306. doi:10.1016/j.jfineco.2019.09.006

9. Zhang H, Chen Y, Rong W, Wang J, Tan J. Effect of social media rumors on stock market volatility: a case of data mining in China. *Front Phys* (2022) 10:987799. doi:10.3389/fphy.2022.987799

10. Ahmed WMA. Stock market reactions to domestic sentiment: panel CS-ardl evidence. *Res Int Bus Finance* (2020) 54:101240. doi:10.1016/j.ribaf.2020.101240

11. Tiwari AK, Abakah EJA, Bonsu CO, Karikari NK, Hammoudeh S. The effects of public sentiments and feelings on stock market behavior: evidence from Australia. *J Econ Behav Organ* (2022) 193:443–72. doi:10.1016/j.jebo.2021.11.026

12. Tetlock PC. Giving content to investor sentiment: the role of media in the stock market. *J Finance* (2007) 62:1139–68. doi:10.1111/j.1540-6261.2007.01232.x

13. Singal M. Effect of consumer sentiment on hospitality expenditures and stock returns. *Int J Hosp Manag* (2012) 31(2):511–21. doi:10.1016/j.ijhm.2011.07.009

14. Wang X, Han X, Chen Z, Bi Q, Guan S, Zou Y. Multi-scale transition network approaches for nonlinear time series analysis. *Chaos Solitons Fractals* (2022) 159:112026. doi:10.1016/j.chaos.2022.112026

15. Xu X. What are customers commenting on, and how is their satisfaction affected? Examining online reviews in the on-demand food service context. Decis. *Support Syst* (2021) 142:113467. doi:10.1016/j.dss.2020.113467

16. Nilashi M, Fallahpour A, Wong KY, Ghabban F. Customer satisfaction analysis and preference prediction in historic sites through electronic word of mouth. *Neural Comput Appl* (2022) 34(16):13867–81. doi:10.1007/s00521-022-07186-5

17. Kumar A, Chakraborty S, Bala PK. Text mining approach to explore determinants of grocery mobile app satisfaction using online customer reviews. *J Retail Consum Serv* (2023) 73:103363. doi:10.1016/j.jretconser.2023.103363

18. Minaee S, Kalchbrenner N, Cambria E, Nikzad N, Chenaghlu M, Gao J. Deep learning-based text classification: a comprehensive review. *ACM Comput Surv* (2022) 54(3):1–40. doi:10.1145/3439726

19. Alantari HJ, Currim IS, Deng Y, Singh S. An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *Int J Res Mark* (2022) 39(1):1–19. doi:10.1016/j.ijresmar.2021.10.011

20. Li H, Chen Q, Zhong Z, Gong R, Han G. E-word of mouth sentiment analysis for user behavior studies. *Inf Process Manag* (2022) 59(1):102784. doi:10.1016/j.ipm.2021.102784

21. Liang D, Dai Z, Wang M. Assessing customer satisfaction of O2O takeaway based on online reviews by integrating fuzzy comprehensive evaluation with AHP and probabilistic linguistic term sets. *Appl Soft Comput* (2021) 98:106847. doi:10.1016/j.asoc.2020.106847

22. Gang Z, Chenglin L. Dynamic measurement and evaluation of hotel customer satisfaction through sentiment analysis on online reviews. *J Organ End User Comput* (2021) 33(6):1–27. doi:10.4018/JOEUC.20211101.oa8

23. Darko AP, Liang D. Modeling customer satisfaction through online reviews: a FlowSort group decision model under probabilistic linguistic settings. *Expert Syst Appl* (2022) 195:116649. doi:10.1016/j.eswa.2022.116649

24. Park J. Combined text-mining/DEA method for measuring level of customer satisfaction from online reviews. *Expert Syst Appl* (2023) 232:120767. doi:10.1016/j.eswa.2023.120767

25. Jiang X, Li SP, Mai Y, Tian T. Study of multinational currency Co-movement and exchange rate stability base on network game. *Finance Res Lett* (2022) 47:102601. doi:10.1016/j.frl.2021.102601

26. Urom C, Ndubuisi G, Guesmi K, Benkraien R. Quantile Co-movement and dependence between energy-focused sectors and artificial intelligence. *Technol Forecast Soc Change* (2022) 183:121842. doi:10.1016/j.techfore.2022.121842

27. Farid S, Karim S, Naeem MA, Nepal R, Jamasb T. Co-movement between dirty and clean energy: a time-frequency perspective. *Energy Econ* (2023) 119:106565. doi:10.1016/j.eneco.2023.106565

28. Wu K, Zhu J, Xu M, Yang L. Can crude oil drive the Co-movement in the international stock market? Evidence from partial wavelet coherence analysis. NORTH Am. J. Econ. *Finance* (2020) 53:101194. doi:10.1016/j.najef.2020.101194

29. Yu X, Xiao K, Liu J. Dynamic Co-movements of COVID-19 pandemic anxieties and stock market returns. *Finance Res Lett* (2022) 46:102219. doi:10.1016/j.frl.2021.102219

30. Wadud S, Gronwald M, Durand RB, Lee S. Co-movement between commodity and equity markets revisited-an application of the Thick pen method. *Int Rev Financ Anal* (2023) 87:102568. doi:10.1016/j.irfa.2023.102568

31. Liu H, Zhao T, Wang S, Li X. A stock rank prediction method combining industry attributes and price data of stocks. *Inf Process Manag* (2023) 60(4):103358. doi:10.1016/j.ipm.2023.103358

32. Zhu C, Fan R, Sun J, Luo M, Zhang Y. Exploring the fluctuant transmission characteristics of air quality Index based on time series network model. *Ecol Indic* (2020) 108:105681. doi:10.1016/j.ecolind.2019.105681

33. Li H, Liu Z. Multivariate time series clustering based on complex network. *Pattern Recognit* (2021) 115:107919. doi:10.1016/j.patcog.2021.107919

34. Fatima S, Uddin M. On the forecasting of multivariate financial time series using hybridization of DCC-GARCH model and multivariate ANNs. *Neural Comput Appl* (2022) 34(24):21911–25. doi:10.1007/s00521-022-07631-5

35. Tang K, Guo T, Shao F, Ma Y, Khattak AJ. Spatial-temporal traffic performance collaborative forecast in urban road network based on dynamic factor model. *Expert Syst Appl* (2023) 225:120090. doi:10.1016/j.eswa.2023.120090

36. Zhou Y, Xie C, Wang GJ, Zhu Y, Uddin GS. Analysing and forecasting co-movement between innovative and traditional financial assets based on complex network and machine learning. *Int Bus Finance* (2023) 64:101846. doi:10.1016/j.ribaf.2022.101846

37. Liu M, Lee CC. Capturing the dynamics of the China crude oil futures: markov switching, co-movement, and volatility forecasting. *Energ Econ* (2021) 103:105622. doi:10.1016/j.eneco.2021.105622

38. Laborda R, Olmo J. Volatility spillover between economic sectors in financial crisis prediction: evidence spanning the great financial crisis and covid-19 pandemic. *Res Int Bus Finance* (2021) 57:101402. doi:10.1016/j.ribaf.2021.101402

39. Men B, Long R, Li Y, Liu H, Tian W, Wu Z. Combined forecasting of rainfall based on fuzzy clustering and cross entropy. *Entropy* (2017) 19(12):694. doi:10.3390/e19120694

40. Wu Y, Qian C, Huang H. Enhanced air quality prediction using a coupled DVMD informer-CNN-LSTM model optimized with dung beetle algorithm. *Entropy* (2024) 26(7):534. doi:10.3390/e26070534

41. Zhang Z, Sun F, Li Q, Wang W, Hu D, Li S. Establishment of the predicting models of the dyeing effect in supercritical carbon dioxide based on the generalized regression neural network and Back propagation neural *network. Processes. Processes (Basel)* (2020) 8(12):1631. doi:10.3390/pr8121631

42. Gul H, Amin A, Adnan A, Huang K. A systematic analysis of link prediction in complex network. *IEEE Access* (2021) 9:20531–41. doi:10.1109/ACCESS.2021.3053995

43. De Souza Moreira FR, Neto Verri FA, Yoneyama T. Maximum visibility: a novel approach for time series forecasting based on complex network theory. *IEEE Access* (2022) 10:8960–73. doi:10.1109/ACCESS.2022.3143106

44. Awijen H, Ben Ameur H, Ftiti Z, Louhichi W. Forecasting oil price in times of crisis: a new evidence from machine learning versus deep learning models. *Ann Oper Res* (2023). doi:10.1007/s10479-023-05400-8

45. Nasiri E, Berahmand K, Li Y. A new link prediction in multiplex networks using topologically biased random walks. *Chaos Solitons Fractals* (2021) 151:111230. doi:10.1016/j.chaos.2021.111230

46. Wen T, Chen H, Cheong KH. Visibility graph for time series prediction and image classification: a review. *Nonlinear Dyn* (2022) 110(4):2979–99. doi:10.1007/s11071-022-08002-4

47. Hao J, Li Z, Wu Z, Ma J. Role of degree and weighted coreness based on endpoints in link prediction. *Front Phys* (2022) 10:1016535. doi:10.3389/fphy.2022.1016535

48. Xu G, Yu Z, Yao H, Li F, Meng Y, Wu X. Chinese text sentiment analysis based on extended sentiment dictionary. *IEEE access* (2019) 7:43749–62. doi:10.1109/access.2019.2907772

49. Wang M, Tian L. From time series to complex networks: the phase space coarse graining. *Phys Stat Mech Its Appl* (2016) 461:456–68. doi:10.1016/j.physa.2016.06.028

50. Kapur JN, Kesavan HK. Entropy optimization principles and their applications. In: VP Singh, M Fiorentino, editors. *Entropy and energy dissipation in water resources*. Dordrecht: Springer Netherlands (1992). p. 3–20. doi:10.1007/978-94-011-2430-0_1

51. Rodriguez-Rodriguez N, Miramontes O. Shannon entropy: an econophysical approach to cryptocurrency portfolios. *Entropy* (2022) 24(11):1583. doi:10.3390/e24111583

52. Su H, Chen D, Pan GJ, Zeng Z. Identification of network topology variations based on spectral entropy. *IEEE Trans Cybern* (2022) 52(10):10468–78. doi:10.1109/TCYB.2021.3070080

53. Hao B, Kovács IA. A positive statistical benchmark to assess network agreement. *Nat Commun* (2023) 14(1):2988. doi:10.1038/s41467-023-38625-z

54. Zhu S, Shang R, Yuan B, Zhang W, Li W, Li Y, et al. DynamicKD: an effective knowledge distillation via dynamic entropy correction-based distillation for gap optimizing. *Pattern Recognit* (2024) 153:110545. doi:10.1016/j.patcog.2024.110545

55. Gupta R, Drzazga-Szczęśniak EA, Kais S, Szczęśniak D. Entropy corrected geometric brownian motion. *Sci Rep* (2024) 14(1):28384. doi:10.1038/s41598-024-79714-3

56. Kukreti V, Pharasi HK, Gupta P, Kumar S. A perspective on correlation-based financial networks and entropy measures. *Front Phys* (2020) 8:553825. doi:10.3389/fphy.2020.00323

57. Wang G, Zha Y, Wu T, Qiu J, Peng J, Xu G. Cross entropy optimization based on decomposition for multi-objective economic emission dispatch considering renewable energy generation uncertainties. *Energy* (2020) 193:116790. doi:10.1016/j.energy.2019.116790

58. Angelidis G, Bratsas C, Makris G, Ioannidis E, Varsakelis NC, Antoniou IE. Global value chains of COVID-19 materials: a weighted directed network analysis. *Mathematics* (2021) 9(24):3202. doi:10.3390/math9243202

59. Wang M, Tian L, Zhou P. A novel approach for oil price forecasting based on data fluctuation network. *Energ Econ* (2018) 71:201–12. doi:10.1016/j.eneco.2018.02.021

60. Mohammadzadeh A, Mirjalili S. Eel and grouper optimizer: a nature-inspired optimization algorithm. *Clust Comput* (2024) 27(9):12745–86. doi:10.1007/s10586-024-04545-w

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 18 March 2025
DOI 10.3389/fphy.2025.1553788

Check for updates

OPEN ACCESS

EDITED BY
Brian Lucey,
Trinity College Dublin, Ireland

REVIEWED BY
Amar Oukil,
Sultan Qaboos University, Oman
Feng Li,
Southwestern University of Finance and Economics, China
Junfei Chu,
Central South University, China

*CORRESPONDENCE
Mingqian Xu,
2022212766@mail.hfut.edu.cn

RECEIVED 31 December 2024
ACCEPTED 14 February 2025
PUBLISHED 18 March 2025

CITATION
Ma Y and Xu M (2025) Financial crisis impact on the efficiency evaluation of university research achievements transformation: a study based on fixed and interactive network data envelopment analysis methods.
Front. Phys. 13:1553788.
doi: 10.3389/fphy.2025.1553788

COPYRIGHT
© 2025 Ma and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Financial crisis impact on the efficiency evaluation of university research achievements transformation: a study based on fixed and interactive network data envelopment analysis methods

Yuping Ma[1,2] and Mingqian Xu[3]*

[1]Anhui Sanxiang Technology Consulting Co., Ltd., Hefei, China, [2]School of Economics and Management, Anhui Jianzhu University, Hefei, China, [3]School of Economics, Hefei University of Technology, Hefei, China

In the context of the financial crisis, this paper introduces an innovative approach to Data Envelopment Analysis (DEA) that is grounded in collaborative theory, aiming to assess the impact of financial factors, particularly fiscal allocations, on the efficiency of transforming research achievements into practical applications within Chinese universities. Based on this methodological framework, the paper constructs an interactive network framework that integrates government, industry, and academic institutions, conceptualizing research activities as a multi-agent, multi-stage complex system. Through an empirical analysis of 79 Chinese universities, we investigate the role of government fiscal grants as a key input factor in influencing the efficiency of research achievement transformation within this system. The findings reveal that strategic allocation of fiscal grants significantly enhances the efficiency of research application, while substantial variations in the efficiency of research achievement transformation exist across different universities. This study further elucidates the intrinsic link between fiscal allocations and the commercialization efficiency of research achievements, providing policymakers with critical insights into the effective distribution of financial resources to facilitate the transformation of research achievements into practical applications. This research not only enriches the application of complex systems theory in higher education but also offers a novel perspective on the role of financial support policies in the commercialization of scientific and technological achievements.

KEYWORDS

financial crisis, university performance evaluation, data envelopment analysis, synergy theory, complex systems

1 Introduction

Set against the backdrop of a financial crisis, this paper introduces a new Data Envelopment Analysis (DEA) methodology, rooted in collaboration theory, to assess the

financial crisis's impact on the efficiency of transforming research achievements in Chinese universities into practical applications. Within this complex system, the collaborative network of universities, enterprises, and government is crucial for propelling significant advancements in science and technology, with fiscal elements, particularly fiscal allocations, playing a pivotal role in directly influencing the direction, scale, and outcomes of scientific research.

Collaborative innovation emerges as an advanced organizational paradigm that leverages the collective strength of these entities to catalyze major scientific and technological innovations, with value-added knowledge at its core and achieving extensive integration. The application of synergy theory is indispensable for enhancing the efficiency of scientific research innovation and achievement transformation. This theory focuses on the common characteristics and synergistic mechanisms among different systems, emphasizing the transition from disorder to order through the interaction of various systems.

Synergy theory comprises three principal elements: synergistic effect, servo principle, and self-organization. The synergistic effect, resulting from the interaction of different systems within an open system, is the internal driving force for the formation of organized structures. The servo principle dictates that fast variables are governed by slow variables and order parameters, which dominate system behavior and elaborate the process of system self-organization. Self-organization allows a system to automatically form certain structures or functions among its internal subsystems according to specific rules.

Given the characteristics of research universities, especially those directly under the administration of the Ministry of Education in China, evaluating their efficiency is crucial for advancing the construction of world-class universities, optimizing resource allocation, and fostering collaborative innovation among universities, enterprises, and government. Efficiency evaluation provides organizations with a potent quantitative tool to inform managerial decisions, with analysis results offering a scientific basis for enhancing organizational efficiency.

As a non-parametric approach, Data Envelopment Analysis (DEA) excels in assessing the efficiency of multi-input, multi-output systems. The fundamental principle of DEA is to compare the actual input-output levels of evaluated decision-making units (DMUs) with the ideal levels that could be improved, with this ratio serving as the efficiency measure of the evaluated units.

However, the assumption of independence among decision units is often violated in real-world scenarios due to mutual interference and interconnected production processes among decision units. Under non-shared fixed and input constraints, changes in the output level of one decision-making unit can affect others, altering the position of the efficient frontier and the desired input-output levels that could be enhanced. Thus, reassessing the comparative performance metrics of parallel interactive decision units with fixed and input constraints presents an efficiency evaluation challenge that demands reconceptualization.

Drawing on the connotations of synergy theory, this paper proposes a novel DEA efficiency evaluation method and applies it to assess the scientific research performance of China's "Double First-Class" universities amidst financial crisis by the Refs. [1, 2], with particular emphasis on the role of fiscal grants in shaping research outcomes and efficiency. We measure the relative efficiency of research universities by proposing a parallel interactive network Data Envelopment Analysis (DEA) model with non-shared fixed inputs. The interactivity is reflected in the division of the university's research activities into research development subsystems and research application subsystems, which interact with each other within the model. The outputs of the research development subsystems, such as papers and monographs, serve simultaneously as inputs for the research application subsystems for the transformation of research results. Conversely, the research funding obtained through enterprises by the research application subsystems is used as input for the research development subsystems. The fixed nature is represented by government funding, which is consumed as a fixed input solely by the research development subsystems. This approach allows us to capture the complex interdependencies and resource allocation within research universities, providing a more refined perspective on the efficiency analysis of higher education research.

The study reveals that under the influence of the 2015 Chinese stock market financial crisis, the strategic allocation of fiscal grants to universities significantly enhanced the efficiency of research outcome application during 2016–2017, offering policymakers critical insights into the effective distribution of financial resources to catalyze the transformation of research outcomes into practical applications. This research not only enriches the application of complex systems theory in higher education but also provides a novel perspective on the role of financial support policies in the commercialization of scientific and technological achievements within the context of financial crisis.

2 Literature review

In recent years, the synergistic theory has garnered significant scholarly interest. Synergistic innovation, a pivotal outcome of this theory, redirects research focus towards the domain of collaborative technological innovation. The concept was initially introduced by Peter Gero from the Sloan Center at the Massachusetts Institute of Technology (MIT) in the U.S., who defined Synergistic innovation as the process by which a network of self-motivated individuals develops a shared vision, communicates ideas, information, and work progress, and collaboratively pursues a common objective [3]. supposed that the group members form a common vision and exchange ideas, information and work status through the network, working together to achieve common goals. Considering the current situation of innovation in China [3], redefined collaborative innovation as a way of guiding by national directional policies and mechanisms, enabling enterprises, universities and other innovation subjects to utilize their respective advantageous capabilities, readjust complementary

Abbreviations: DEA, Data Envelopment Analysis; DMU, Decision-Making Unit; GGF, Government Grant Funds; FSODEA, Fixed-Sum Output DEA; EEFDEA, Equilibrium Efficiency Frontier DEA; GEEFDEA, Generalized Equilibrium Efficiency Frontier DEA; R&D, Research and Development; FS, Number of teachers; GGF, Government Grant Funds; RF, Research funding; AS, Awards; PM, Publication of monographs; PP, Published papers; TNS, Number of Natural Science Funds; TTI, Income from technology transfer.

resources, thus realizing complementary advantages, promoting the diffusion of innovative technologies and industrialization, and accelerating techno-logical innovation and the industrialization of technological achievements. The industrialization of technological innovation and scientific and technological achievements will be accelerated.

This paper posits that synergy theory enhances innovation and practicality during the research and development phase by fostering interdisciplinary collaboration. Specifically, research teams from different disciplines can integrate their respective knowledge and technological strengths through synergistic innovation, forming comprehensive research plans that provide a broader perspective for the generation of scientific achievements by the Ref. [3]. In the stage of application and transformation of research outcomes, synergy theory facilitates the close alignment of scientific achievements with market demands through collaboration between enterprises and universities. Considering the three major theoretical frameworks of synergy theory, the synergistic effect enables parties to pool their advantageous resources through resource sharing and information exchange, thereby improving the efficiency and quality of outcome transformation by the Ref. [4]. The servo principle guides research institutions to adjust their research directions and strategies in a timely manner according to market and policy changes, adapting to the demands of the external environment. The self-organization capability allows research teams to autonomously manage and optimize their internal operations, forming efficient organizational structures and operational mechanisms, thus enhancing the output and application efficiency of scientific achievements.

In academia, the traditional methodology for evaluating the efficiency of multi-input, multi-output decision-making units (MIMDU), known as Data Envelopment Analysis (DEA), has attracted significant global scholarly attention. Seminal research in this domain employed the standard DEA to evaluate the performance of higher education institutions, notably by quantifying the relative efficiency of 79 Chinese universities from 2003 to 2004 [5]. classified the departments were methodically sorted into quartets for the sake of cluster analysis and the breakdown of efficiency, taking into account their diverse characteristics. This approach diverges from the straightforward production framework depicted by the DEA model, which does not delve into the same granularity of features. Taking into account qualitative aspects within the DEA approach, Ref. [6] assessed the issue of fuzzy efficiency to aid in decision-making processes. Furthermore, Ref. [7] introduced an enhanced interval DEA model, which addressed the challenge of zero inputs when evaluating the efficiency of certain DMUs, and this model was subsequently applied to Iranian academic institutions. Similarly [8], evaluated the efficiency problem of private universities with the help of DEA modeling. The aforementioned studies indicate that the conventional DEA model, characterized by a "black box" design, has seen extensive application in the assessment of university efficiency. However, contemporary scholarly works frequently portray the production process of universities as a simple construct consisting only of initial inputs and final outputs. This rudimentary black box model is inadequate for capturing the complexity of actual production processes, potentially omitting crucial details within these activities. As a result, the development of the network DEA approach has emerged to overcome the limitations inherent in traditional DEA models, particularly in the context of university efficiency assessments.

[4] introduced a network DEA methodology with the objective of unveiling the "black box" of the production process, typically perceived as an aggregation of various subsystems. The network DEA methodology aims to reveal the internal workings of the production process by viewing it as an integrated system of distinct subsystems. By dissecting the internal production mechanisms, this approach enables a more comprehensive understanding of an organization's efficiency, linking the overall performance directly to the efficacy of its individual subsystems. In today's context, universities fulfill various critical roles, including teaching, research, and technology transfer, with each role representing a key subsystem that significantly impacts the institution's overall efficiency.

The network DEA approach has been increasingly utilized to assess university efficiency by modeling the internal production processes, thereby providing new perspectives and broadening the scope of efficiency assessment research within the academic community [9]. assessed the educational, scholarly, and financial efficiency of Taiwanese universities, with the educational and scholarly stages incorporating outputs from the financial efficiency stage. The findings indicated that universities excelled in the cost efficiency phase [10]. introduced a two-stage network configuration DEA methodology for assessing the efficiency of travel and recreation departments across 34 universities [11]. developed a network DEA model that puts into perspective parallel interdependent systems and utilized it to evaluate the productivity of universities. Associated with the "985 project." [12] employed a semidefinite programming method to compute a generalizable two-stage model, which was then applied to R&D activities within China. Concurrently, the efficiency of China's regional R&D processes was also evaluated [13]. investigated the use of the Luenberger productivity indicator to gauge efficiency dynamics in Chinese universities. Notably, the research is grounded in a network framework that incorporates the element of time [14]. developed a multi-period DEA approach that accounts for feedback mechanisms, which was used to classify sociological inquiry into development and innovation phases [15]. introduced an additive DEA model to assess the efficiency of 38 academic departments within Chinese universities, considering the heterogeneity among university faculties [16]. employed an augmented two-step network configuration to appraise the efficiency of 52 Chinese universities, culminating in the finding that approximately two-thirds of these institutions are deemed inefficient [17]. built a decentralized and centralized model that includes networked institutions for the purpose of evaluating the efficiency of elite universities in China. In order to address the deficiencies of static systems [18], conceptualized university R&D activities as an intertemporal production process and constructed a parallel DEA model that captures the intertemporal dynamics. Considering the fairness between the leadership step and further steps [19], constructed a three-stage network with a tandem structure and applied it to the R&D innovation of China's hightech industry. Echoing this trend, the network DEA approach has gained prominence due to its benefits in assessing organizational performance and has been extensively utilized to evaluate the efficiency of various entities, including universities and industrial firms. For example [20], used a two-stage network DEA model to measure energy efficiency

and explored the differences with full efficiency [21]. advanced a progressive DEA model that surpasses earlier methods in detecting alterations in system efficiency.

Scholars, in their quest to refine the evaluation of production structures through various network DEA methods, have increasingly directed their focus towards the critical issues of input and output fixation and the interplay between decision-making units (DMUs). The oversight of fixation and the characteristics of inputs and outputs between DMUs can lead to imprecise efficiency valuations. This realization underscores the necessity for a nuanced approach that accounts for these intricacies, thereby enhancing the accuracy and reliability of efficiency assessments within academic research. In view of this [22], first investigated the problem of evaluating participating countries with a fixed total number of medals in Olympic competitions and proposed a zero sum gains (ZERO SUM Gains) DEA model. Subsequently [23], extended the zero-sum gains DEA model to consider the case of non-desired outputs and assessing the efficiency of carbon dioxide emissions among the parties to the Kyoto Protocol [24]. established a two-stage efficiency evaluation model with substage efficiency decomposition model that satisfies the shared output fixation and constraints, and applied it to the problem of evaluating the energy conversion and utilization efficiency in 30 provincial-level regions in China [25]. considering that government grants are consumed by two subsystems, scientific and technological research and results transformation, a parallel interaction network DEA method based on shared fixed inputs was established and evaluated the efficiency of 58 Chinese research universities. However, this paper argues that not all decision-making units have shared fixed sum constraints. Take universities, for example,. When measuring their performance, government grants are mostly used in scientific and technological research in reality, and the source of funding for the transformation of achievements does not include government grants, so the shared fixed-sum input condition does not hold.

Financial crises can significantly impact the efficiency of the complex production systems of university research activities, particularly the stage of transforming research outcomes into practical applications [1]. focused on the efficiency of Italian and German universities in converting public funding into multiple university outputs (i.e., graduating students, publishing research, and patenting activities) following the 2008 financial crisis [2]. considered the constraints on public funding after the economic crisis and empirically analyzed the impact of the economic crisis on the technical efficiency of research and development activities in state universities [26]. analyzed the significant impact of the financial crisis on various stakeholders in higher education activities, suggesting that the crisis may lead to reduced funding for education from governments, private sectors, and households [27]. explored the determinants of cost efficiency in public universities during the economic crisis, systematically quantifying the efficiency impact factors on universities [28]. considered the differences in the impact of the global financial crisis on private and public universities, indicating that reliance on public and endowment funds can significantly enhance the efficiency of university activities after the crisis. Unlike previous studies, this paper considers the actual context of Chinese universities, using government financial subsidies as a key input factor to assess the impact of financial elements (especially fiscal allocation) on the efficiency of Chinese universities in transforming research outcomes into practical applications.

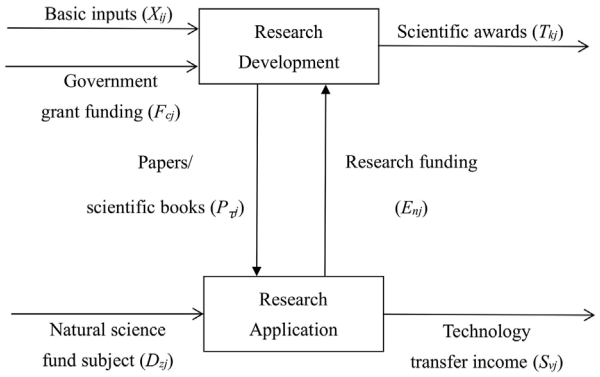


FIGURE 1
Structure of the two systems of research and development - results transformation and application.

Thus, this study tackles the two-stage efficiency evaluation challenge within a networked system, constrained by specific fixed inputs and limitations, and is rooted in synergy theory. Initially, the paper establishes an equilibrium efficiency frontier that adheres to the non-shared fixed and input constraints. Subsequently, it develops a two-system efficiency evaluation model predicated on this equilibrium efficient frontier. Recognizing that the efficiency evaluation of subsystems is influenced by various factors, including parameter settings and subjective roles, which can yield non-unique results, the paper addresses the holistic efficiency evaluation of the two systems. Ultimately, the study employs empirical data from 79 Chinese research universities to validate the methodology.

3 Methodological models

3.1 Conventional concurrent interactive DEA network model with communal inputs

Consider homogeneous decision-making units with a two-system production structure as shown in Figure 1. In the first subsystem, each DMU_j uses m basic inputs X_{ij} (i = 1,2, …m), F_{cj} (c = 1,2, …,d) and E_{nj} (n = 1,2, …,q) obtained from the second subsystem to produce $P_{\tau j}$ (τ = 1,2, …, o) and T_{kj} (k = 1,2, …, q). In the second subsystem, each DMU_j uses $P_{\tau j}$ (τ = 1,2, …, w) obtained from the first subsystem and D_{zj} (z = 1,2, …, d) to produce S_{vj} (v = 1,2, …,o) and E_{nj} (n = 1,2, …, l). It is important to note that the E_{n_j} and $P_{\tau j}$ are interaction indicators, and F_{cj} are the fixed inputs of the first subsystem. Thus, the two subsystems are interwoven rather than functioning in isolation from one another.

Following the approach of previous studies [29], adopted the variable returns to scale (VRS) hypothesis, positing that for a given Decision-Making Unit (DMU_0), the overall system efficiency is calculated as a composite of the individual efficiencies of the two subsystems, weighted accordingly. This concept is articulated in model (Equation 1), which encapsulates the synthesis of subsystem efficiencies into a comprehensive measure of performance.

$$Max E_0 = \varphi_1 * E_{10} + \varphi_2 * E_{20}$$

$$s.t. E_{1j} = \frac{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{k=1}^{w} \mu_k T_{kj} + u_1}{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj}} \le 1$$

$$E_{2j} = \frac{\sum_{n=1}^{o} \phi_n E_{nj} + \sum_{q=1}^{l} \beta_q S_{qj} + u_2}{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}} \le 1 \quad (1)$$

$$v_i, \delta_c, \phi_n, \sigma_\tau, y_z, \mu_k, \beta_q \ge \varepsilon, \forall i, c, n, z, \tau, k, q$$

u_1 u_2 free in sign

Among them E_j, E_{1j} and E_{2j} represents the overall efficiency, first subsystem efficiency, and second subsystem efficiency, respectively. $v_i, \delta_{1c}, \delta_{2c}, \phi_n, \sigma_\tau, y_z, \mu_k$ and β_q are unknown positive multipliers associated with the input and output variables. Specifically [30], proposed when $\mu_1 = \mu_2 = 0$ holds, the model is constructed based on the assumption of constant returns to scale (CRS). To achieve a viable solution [31, 32], supposed the intermediate measures E_{nj} and $P_{\tau j}$ must be assigned equal multipliers across the subsystems.

Based on the theory of synergy, φ_1 and φ_2 are the weights of the first and second subsystems, indicating the relative neediness of each subsystem. A practical approach to determining these weights is to compute the input consumption proportions for the two subsystems. Since $\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj}$ and $\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}$ represent the inputs of the two subsystems, respectively $\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj} + \sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}$ is the total input of the first subsystem. Therefore φ_1 and φ_2 are assigned as follows, the $\varphi_1 + \varphi_2 = 1$.

$$\varphi_1 = \frac{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj}}{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj} + \sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}}$$

$$\varphi_2 = \frac{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}}{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj} + \sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}} \quad (2)$$

$\varphi_1 \ge a$, $\varphi_2 \ge b$, a and b are the given minimum levels.

In model (Equation 2), it is required that φ_1 and φ_2 satisfy some conditions to prevent distorted and inconceivable outcomes. Therefore, we require $\varphi_1 \ge a, \varphi_2 \ge b$, where a and b are the given minimum levels. Integrating model (Equation 2) into the preceding model (Equation 1) allows us to deduce the subsequent model (Equation 3), which is used to compute the overall efficiency.

$$Max E_0 = \frac{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau 0} + \sum_{k=1}^{w} \mu_k T_{k0} + u_1 + \sum_{n=1}^{o} \phi_n E_{n0} + \sum_{q=1}^{l} \beta_q S_{q0} + u_2}{\sum_{i=1}^{m} v_i X_{i0} + \sum_{c=1}^{d} \delta_c F_{c0} + \sum_{n=1}^{o} \phi_n E_{n0} + \sum_{\tau=1}^{q} \sigma_\tau P_{\tau 0} + \sum_{z=1}^{h} y_z D_{z0}}$$

$$s.t. E_{1j} = \frac{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{k=1}^{w} \mu_k T_{kj} + u_1}{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj}} \le 1$$

$$E_{2j} = \frac{\sum_{n=1}^{o} \phi_n E_{nj} + \sum_{q=1}^{l} \beta_q S_{qj} + u_2}{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}} \le 1$$

$$\frac{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj}}{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj} + \sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}} \ge a \quad (3)$$

$$\frac{\sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}}{\sum_{i=1}^{m} v_i X_{ij} + \sum_{c=1}^{d} \delta_c F_{cj} + \sum_{n=1}^{o} \phi_n E_{nj} + \sum_{\tau=1}^{q} \sigma_\tau P_{\tau j} + \sum_{z=1}^{h} y_z D_{zj}} \ge b$$

$$v_i, \delta_c, \phi_n, \sigma_\tau, y_z, \mu_k, \beta_q \ge \varepsilon, \forall i, c, n, z, \tau, k, q$$

u_1 u_2 *free in sign*

Model (Equation 3) adheres to the condition that the weights allocated to each subsystem exceed a predetermined threshold, with the resultant efficiency E_0 of DMU_0 being characterized as the system's overall efficiency. This metric typically varies within the interval from 0 to 1. A value of E_0 = 1, signifies that DMU_0 is deemed efficient. It should be highlighted that Model (Equation 3) is inherently nonlinear, necessitating the investigation of its linear counterpart, details of which will be delineated subsequently.

3.2 Concurrent interacting DEA network architecture incorporating stationary and input limitations within the realm of synergy theory

[33, 34] proposed the preceding DEA model of an integrated interconnected network with traditional shared inputs is particularly noteworthy: for one thing, the quantity of shared inputs in individual subsystems needs to be accurately measured, but based on the theory of synergy, in many production and life practices [35-37], only a total amount of shared inputs can be observed F_{cj} and it is not possible to accurately measure the quantity of individual subsystems.The second is that the total amount of shared inputs is fixed. Secondly, the total amount of shared input is fixed and consumed by two subsystems at the same time, there is a shared fixed and input constraint, but based on the research of synergy theory, we find that in the real production and life practice there often exists a non-shared fixed and input constraint, the total amount of inputs of all the decision-making units is determined, and they influence each other in the production process. For example, in the university production process, government grants are only used as inputs for scientific research and are not used by outcome transformation, but scientific research and outcome transformation influence each other through papers, monographs and research funding. When it is observed that F_{cj} rather than $F_{c_j}^1$ and $F_{c_j}^2$, and considering non-shared fixed and input constraints, it is difficult for model (Equation 3) to accurately appraise the relative efficiencies of decision-making entities with a two-system production structure as shown in Figure 1.

For efficiency evaluation problems with fixed-sum input constraints, changes in the efficient frontier must be taken into account when calculating the relative efficiency because the production process still does not satisfy the independence requirement [38, 39]. introduced a framework that utilizes the equilibrium efficiency horizon in the black-box scenario of decision units to find such a common efficient frontier, which can make all decision units lie on the efficient frontier at the same time when output fixing and constraints are satisfied.

In assessing the efficiency of Decision-Making Units (DMUs), the conventional DEA approach establishes the optimal efficiency frontier using efficient DMUs. Nonetheless, it overlooks the

scenario where certain DMUs are subject to fixed sum constraints. Consequently, leveraging the minimum adjustment quantity strategy [40], introduced the FSODEA (Fixed-Sum Output DEA) model, which puts into perspective the fixed-sum output constraint. In this FSODEA framework, the DMUs under evaluation are deemed efficient following the adjustment of outputs for other DMUs, with the optimal efficiency frontiers differing according to the DMU in question. Following this [41], crafted the EEFDEA (Equilibrium Efficiency Frontier DEA) model, capable of assessing all DMUs against a shared equilibrium efficiency frontier. However, EEFDEA is not without its limitations, such as the considerable computational complexity and the dependence of evaluation results on the pre-established "order" of DMUs. Building upon these [38], developed the GEEFDEA (Generalized Equilibrium Efficiency Frontier DEA) model, which retains the benefits of both FSODEA and EEFDEA and establishes a generalized equilibrium efficiency frontier. To evaluate parallel research activities with non-shared fixed and variable inputs, we constructed the parallel GEEFDEA network framework. To assess the efficiency of two systems with non-shared fixed and input constraints, we have developed model (Equation 4) to ascertain the common EEF.

$$Min\sum_{j=1}^{n}\sum_{c=1}^{d}\delta_c w_{cj}$$

$$s.t.\frac{\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{k=1}^{w}\mu_k T_{kj}+u_1+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{q=1}^{l}\beta_q S_{qj}+u_2}{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c\left(F_{cj}+\pi_{cj}\right)+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}=1$$

$$\sum_{j=1}^{n}\pi_{cj}=0$$

$$w_{cj}=Max\left\{0,\pi_{cj}\right\}$$

$$F_{cj}+\pi_{cj}\geq 0 \tag{4}$$

$$\frac{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}}{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}\geq a$$

$$\frac{\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}\geq b$$

$$v_i,\delta_c,\phi_n,\sigma_\tau,y_z,\mu_k,\beta_q\geq\varepsilon,\forall i,c,n,z,\tau,k,q$$

u_1 u_2 *free in sign*

Where π_{cj} is the GGF adjustment value for the first subsystem. It is the level of adjustment fixed and invested in each DMU to achieve the common EEF. In addition, the π_{cj} is symbol free. In model (Equation 4), the first restrictive factor is that all DMUs reach the effective boundary at the same time following the tuning of input parameters. The second restrictive factor meets the fixed and input stipulations. The fourth restrictive factor $w_{cj}=\mathrm{Max}\left\{0,\pi_{cj}\right\}$ indicates that w_{cj} in the case of π_{cj} non-negative is equal to π_{cj}.The fifth and sixth constraints guarantee that the adjusted GGF is on the positive side. Concerning the seventh and eighth limitations, as in model (Equation 3), the weights of the two subsystems must be within a predetermined interval. Since the model (Equation 4) is $w_{cj}=\mathrm{Max}\left\{0,\pi_{cj}\right\}$ difficult to solve, model (Equation 5) is proposed.

$$Min\sum_{j=1}^{n}\sum_{c=1}^{d}\delta_c\left|\pi_{cj}\right|$$

$$s.t.\frac{\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{k=1}^{w}\mu_k T_{kj}+u_1+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{q=1}^{l}\beta_q S_{qj}+u_2}{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c\left(F_{cj}+\pi_{cj}\right)+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}=1$$

$$\sum_{j=1}^{n}\pi_{cj}=0$$

$$F_{cj}+\pi_{cj}\geq 0$$

$$\frac{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}}{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}\geq a \tag{5}$$

$$\frac{\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}{\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}}\geq b$$

$$v_i,\delta_c,\phi_n,\sigma_\tau,y_z,\mu_k,\beta_q\geq\varepsilon,\forall i,c,n,z,\tau,k,q$$

u_1 u_2 free in sign

Since model (Equation 5) is still nonlinear, we developed a two-phase strategy to transform model (Equation 5) into model (Equation 6). In the first step, we let $\delta_c\pi_{cj}=\pi'_{cj}$, and then we obtained model (Equation 6).

$$Min\sum_{j=1}^{n}\sum_{c=1}^{d}\left|\pi'_{cj}\right|$$

$$\begin{aligned}&\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{k=1}^{w}\mu_k T_{kj}+u_1+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{q=1}^{l}\beta_q S_{qj}+u_2\\&=\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{c=1}^{d}\pi'_{cj}\\&+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}\end{aligned}$$

$$\sum_{j=1}^{n}\pi'_{cj}=0$$

$$\delta_c F_{cj}+\pi'_{cj}\geq 0 \tag{6}$$

$$\begin{aligned}&\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{c=1}^{d}\pi'_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}\\&+\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}\geq H\end{aligned}$$

$$\begin{aligned}&a\left(\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau\right.\\&\left.P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}\right)-\left(\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}\right)\leq 0\end{aligned}$$

$$\begin{aligned}&b\left(\sum_{i=1}^{m}v_i X_{ij}+\sum_{c=1}^{d}\delta_c F_{cj}+\sum_{n=1}^{o}\phi_n E_{nj}+\sum_{\tau=1}^{q}\sigma_\tau\right.\\&\left.P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}\right)-\left(\sum_{\tau=1}^{q}\sigma_\tau P_{\tau j}+\sum_{z=1}^{h}y_z D_{zj}\right)\leq 0\end{aligned}$$

$$v_i,\delta_c,\phi_n,\sigma_\tau,y_z,\mu_k,\beta_q\geq\varepsilon,\forall i,c,n,z,\tau,k,q$$

u_1 u_2 free in sign

H in model (Equation 6) ensures that the model's ratio includes a positive lower fraction and that the efficiency value of the target unit being evaluated is greater than zero. In the second step, we let $\psi_{cj}=\left|\pi'_{cj}\right|+\pi'_{cj}$, $\eta_{cj}=\left|\pi'_{cj}\right|-\pi'_{cj}$, obtain the equivalent

linear model (Equation 7).

$$
\begin{aligned}
&Min\frac{1}{2}\sum_{j=1}^{n}\sum_{c=1}^{d}\left(\psi_{cj}+\eta_{cj}\right)\\
&\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}+\sum_{k=1}^{w}\mu_{k}T_{kj}+u_{1}\\
&\quad+\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{q=1}^{l}\beta_{q}S_{qj}+u_{2}\\
&\quad=\sum_{i=1}^{m}v_{i}X_{ij}+\sum_{c=1}^{d}\delta_{c}F_{cj}+\frac{1}{2}\sum_{c=1}^{d}\left(\psi_{cj}-\eta_{cj}\right)\\
&\quad+\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}+\sum_{z=1}^{h}y_{z}D_{zj}\\
&\sum_{j=1}^{n}\left(\psi_{cj}-\eta_{cj}\right)=0\\
&2\delta_{c}F_{cj}+\left(\psi_{cj}-\eta_{cj}\right)\geq 0
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
&\sum_{i=1}^{m}v_{i}X_{ij}+\sum_{c=1}^{d}\delta_{c}F_{cj}+\frac{1}{2}\sum_{c=1}^{d}\left(\psi_{cj}-\eta_{cj}\right)+\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{\tau=1}^{q}\sigma_{\tau}\\
&\quad P_{\tau j}+\sum_{z=1}^{h}y_{z}D_{zj}\geq H\\
&a\left(\sum_{i=1}^{m}v_{i}X_{ij}+\sum_{c=1}^{d}\delta_{c}F_{cj}+\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{\tau=1}^{q}\sigma_{\tau}\right.\\
&\quad\left.P_{\tau j}+\sum_{z=1}^{h}y_{z}D_{zj}\right)-\left(\sum_{i=1}^{m}v_{i}X_{ij}+\sum_{c=1}^{d}\delta_{c}F_{cj}+\sum_{n=1}^{o}\phi_{n}E_{nj}\right)\leq 0\\
&b\left(\sum_{i=1}^{m}v_{i}X_{ij}+\sum_{c=1}^{d}\delta_{c}F_{cj}+\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{\tau=1}^{q}\sigma_{\tau}\right.\\
&\quad\left.P_{\tau j}+\sum_{z=1}^{h}y_{z}D_{zj}\right)-\left(\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}+\sum_{z=1}^{h}y_{z}D_{zj}\right)\leq 0\\
&v_{i},\delta_{c},\phi_{n},\sigma_{\tau},y_{z},\mu_{k},\beta_{q}\geq\varepsilon,\forall i,c,n,z,\tau,k,q
\end{aligned}
$$

$u_1 u_2$ free in sign

A Data Envelopment Analysis (DEA) model has been developed to characterize a comprehensive equilibrium efficient frontier for network arrangements that interact concurrently, accounting for both fixed and variable inputs. By solving the model (Equation 7), the optimal tuning that enables all DMUs to be simultaneously effective in achieving a common EEF is computed π_{cj}^{*}.

The previous section determines the common EEF, on the basis of which we can evaluate the efficiency of each DMU after obtaining the π_{cj}^{*} the efficiency of each DMU after evaluating the optimal solution. As [39] described, the efficiency obtained from the GEEF model parallels the concept of super-efficiency, as the DMUs under evaluation are omitted from the formation of the common EEF (Equilibrium Efficiency Frontier). Analogously, the efficiency score does not exceed 1 for DMUs located within the common EEF, but it can be greater than 1 for those DMUs that lie outside the common EEF. Model (Equation 8) is employed to calculate the aggregate efficiency of a simultaneous engagement system that operates with non-shared fixed and input constraints.

$$
\begin{aligned}
&MaxE_{0}^{*}=\frac{\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{k=1}^{w}\mu_{k}T_{k0}+u_{1}+\sum_{n=1}^{o}\phi_{n}E_{n0}+\sum_{q=1}^{l}\beta_{q}S_{q0}+u_{2}}{\sum_{i=1}^{m}v_{i}X_{i0}+\sum_{c=1}^{d}\delta_{c}F_{c0}+\sum_{n=1}^{o}\phi_{n}E_{n0}+\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{z=1}^{h}y_{z}D_{z0}}\\
&s.t.E_{1j}=\frac{\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}+\sum_{k=1}^{w}\mu_{k}T_{kj}+u_{1}}{\sum_{i=1}^{m}v_{i}X_{ij}+\sum_{c=1}^{d}\delta_{c}\left(F_{cj}+\pi_{cj}^{*}\right)+\sum_{n=1}^{o}\phi_{n}E_{nj}}\leq 1\\
&E_{2j}=\frac{\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{q=1}^{l}\beta_{q}S_{qj}+u_{2}}{\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}+\sum_{z=1}^{h}y_{z}D_{zj}}\leq 1\\
&\frac{\sum_{i=1}^{m}v_{i}X_{i0}+\sum_{c=1}^{d}\delta_{c}F_{c0}+\sum_{n=1}^{o}\phi_{n}E_{n0}}{\sum_{i=1}^{m}v_{i}X_{i0}+\sum_{c=1}^{d}\delta_{c}F_{c0}+\sum_{n=1}^{o}\phi_{n}E_{n0}+\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{z=1}^{h}y_{z}D_{z0}}\geq a\\
&\frac{\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{z=1}^{h}y_{z}D_{z0}}{\sum_{i=1}^{m}v_{i}X_{i0}+\sum_{c=1}^{d}\delta_{c}F_{c0}+\sum_{n=1}^{o}\phi_{n}E_{n0}+\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{z=1}^{h}y_{z}D_{z0}}\geq b\\
&v_{i},\delta_{c},\phi_{n},\sigma_{\tau},y_{z},\mu_{k},\beta_{q}\geq\varepsilon,\forall i,c,n,z,\tau,k,q
\end{aligned}
\tag{8}
$$

$u_1 u_2$ *free in sign*

The maximum value of the objective function at optimality of model (Equation 9) will be E_0^* is defined as the value of DMU_0 the total efficiency of the model, which $v_i^*, \delta_c^*, \phi_n^*, \sigma_\tau^*, y_z^*, \mu_k^*, \beta_q^*$ It is the value of the $v_i, \delta_c, \phi_n, \sigma_\tau, y_z, \mu_k, \beta_q$ the best possible resolution of the model. Upon acquiring the overall efficiency from model (Equation 9) E_0^*, the efficiency of all subsystems can be figured out according to the formula of model (Equation 1). Due to the impact of various factors, including parameter settings, the model (Equation 9) does not yield a unique optimal solution for the subsystems, and consequently, the individual efficiency scores for each subsystem are not unique either. As a result, this study focuses exclusively on assessing the aggregate efficiency of the two systems.

$$
\begin{aligned}
&MaxE_{0}^{*}=\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{k=1}^{w}\mu_{k}T_{k0}+u_{1}+\sum_{n=1}^{o}\phi_{n}E_{n0}+\sum_{q=1}^{l}\beta_{q}S_{q0}+u_{2}\\
&s.t.\sum_{i=1}^{m}v_{i}X_{i0}+\sum_{c=1}^{d}\delta_{c}F_{c0}+\sum_{n=1}^{o}\phi_{n}E_{n0}+\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}\\
&\quad+\sum_{z=1}^{h}y_{z}D_{z0}=1\\
&\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}+\sum_{k=1}^{w}\mu_{k}T_{kj}+u_{1}\\
&\quad-\sum_{i=1}^{m}v_{i}X_{ij}-\sum_{c=1}^{d}\delta_{c}\left(F_{cj}+\pi_{cj}^{*}\right)-\sum_{n=1}^{o}\phi_{n}E_{nj}\leq 0\\
&\sum_{n=1}^{o}\phi_{n}E_{nj}+\sum_{q=1}^{l}\beta_{q}S_{qj}+u_{2}-\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau j}-\sum_{z=1}^{h}y_{z}D_{zj}\leq 0\\
&a-\left(\sum_{i=1}^{m}v_{i}X_{i0}+\sum_{c=1}^{d}\delta_{c}F_{c0}+\sum_{n=1}^{o}\phi_{n}E_{n0}\right)\leq 0\\
&b-\left(\sum_{\tau=1}^{q}\sigma_{\tau}P_{\tau 0}+\sum_{z=1}^{h}y_{z}D_{z0}\right)\leq 0\\
&v_{i},\delta_{c},\phi_{n},\sigma_{\tau},y_{z},\mu_{k},\beta_{q}\geq\varepsilon,\forall i,c,n,z,\tau,k,q
\end{aligned}
\tag{9}
$$

$u_1 u_2$ *free in sign*

In addressing the rationality of the model's assumptions, this paper conducts a validation and discussion of the assumptions. Drawing on the theory of organizational interdependence, the interactions and dependencies between organizations influence their behavior and performance. In the context of research and development, the interdependence between universities, enterprises, and the government is particularly evident. For instance, universities provide foundational research outcomes, enterprises transform these outcomes into commercial products, and the

government supports this process through policies and funding. This interdependence means that the decisions and activities of each organization are no longer independent but rather affect and constrain each other. This confirms the model's assumption that decision-making units interfere with each other and that production processes are interdependent.

4 Application validation

As China's higher education sector experiences rapid growth, the evaluation of research efficiency and the commercialization of scholarly achievements in academic institutions have become focal points of interest. This assessment is integral to the national innovation system, capturing the attention of educational policymakers and scholars alike. The effectiveness of converting research outcomes into practical applications is a critical issue for those formulating educational strategies and academic discussions. Within the framework of China's "double first-class" initiative, a scientific evaluation of the efficiency of 79 Chinese research universities, grounded in synergy theory, aids in the optimal allocation of resources and fosters the substantive development of these universities [42, 43]. This, in turn, bolsters the overall competitiveness of the nation's educational landscape. This study aims to offer valuable insights for enhancing the management and educational quality of Chinese research universities and, concurrently, to contribute Chinese insights and solutions to the global discourse on higher education.

4.1 Indicators and data

This section employs the efficiency evaluation of China's research universities as a case study, utilizing the methodology outlined in the previous section to assess the academic efficiency of 79 such institutions. According to the synergy theory, the research process within universities encompasses three key stakeholders: universities, enterprises, and the government. This process can be conceptualized as an integrated network framework comprising a research and development (R&D) subsystem and a subsystem for the transformation and application of research outcomes. A distinctive feature of this framework is the fixed government funding allocated to the R&D subsystem, with outputs from the R&D subsystem—such as theses and monographs—serving as inputs back into the system. Additionally, research grants procured through enterprises are also considered inputs to the R&D subsystem. In this analysis, we focus on scenarios where the total government funding is a predetermined fixed amount. The structure of the two systems, namely, research and development and the translation and application of research results, is illustrated in Figure 1.

Taking into account the availability and recency of data, this study employs data from 79 Chinese research universities in 2016 to evaluate the efficiency of research and development (R&D) as well as the transformation of achievements. All of these institutions are participants in the "211 Project," and they have managed a significant number of complex and demanding research projects. It is important to note that due to data accessibility limitations, a subset of these universities with a research focus was excluded from the study. The data was extracted from the Ministry of Education's Basic Statistics Compendium for Universities and the Education Blue Book 2016–2017. To account for output lags, a 1-year lag period for outputs was applied in the analysis. The model was executed using MATLAB R2022b on a system equipped with an Intel Core i7 CPU, 16 GB of RAM, and Windows 11 operating system.

As shown in Figure 1, in the research development system, each of the DMU_j uses m basic inputs X_{ij} (i = 1,2, …m), government funding (GGF) F_{cj} (c = 1,2, …,d) and research funding from the research results translation application subsystem E_{nj} (n = 1,2, …, o) to produce papers, scientific monographs P_{rj} (n = 1,2, …,o) to generate papers, scientific monographs (nτ = 1,2, …, q) and scientific and technical awards T_{kj} (k = 1,2, …, w). In the research application system for the translation of research results, the papers and scientific monographs obtained from the research and development subsystem are utilized P_{rj} (τ = 1,2, …, q) and natural science fund topics D_{zj} (z = 1,2, …, h) to generate income from technology transfer S_{vj} (v = 1,2, …,l) and research funding E_{nj} (n = 1,2, …, o). It should be noted that the E_{n_j} and P_{rj} are interaction indicators, and F_{cj} are the collaborative inputs of the a duo of subsystems. Consequently, the two subsystems are intricately interconnected, functioning in unison rather than in isolation.

As shown in Figure 1, the core components of the research and development (R&D) process include faculty members and graduate students. The government offers grants, classified as inputs, with a 1-year term. To comprehensively evaluate the R&D subsystem's performance, outputs are categorized as scientific awards, published monographs, and high-quality papers. Scientific awards include both national and provincial honors. High-quality papers are characterized by their presence in Chinese core journals and inclusion in the Science Citation Index. In the translation and application phases of research outcomes, the allocation of natural science funds serves as an input variable with two key functions: it indicates the capacity for scientific discovery transformation and highlights the national emphasis on the relevant field. Furthermore, the success of implementing research findings is largely contingent upon the resources provided by natural science funding, underscoring its critical role in the practical application of research achievements. Citing Refs. [39, 44], research funding and income from technology transfers are identified as the outputs of this activity. Specifically, income from technology transfer refers to the payments received by the university from enterprises or other organizations in exchange for the university's provision of technical services. The selection of indicator data and the units for these indicators, which are essential for evaluating the overall efficiency of the R&D and research results transformation and application systems in this study, are presented in Table 1.

4.2 Results and discussion

This paper selected the data of 79 research universities in China in 2016 for empirical research, and Table 2 lists the descriptive statistics of the relevant indicators, including the mean, median, standard deviation, maximum and minimum values.

TABLE 1 Selection of indicators.

Research process		Variable indicators	Unit (of measure)
Research and development (R&D)	Input metrics	Number of teachers (FS)	a hundred people
		Research and development staff (GS)	a hundred people
		Government Grant (GGF)	million dollars
		Research funding (RF)	million dollars
	Output metrics	Awards (AS)	term (in a mathematical formula)
		Publication of monographs (PM)	troops
		Published papers (PP)	hundred articles
Translation and application of research results	Input metrics	Number of Natural Science Funds (TNS)	decathlon (athletics)
		Publication of monographs (PM)	troops
		Published papers (PP)	hundred articles
	Output metrics	Research funding (RF)	million dollars
		Income from technology transfer (TTI)	hundred thousand dollars

In this paper, the optimal efficiency of the application system for research development and translation of research results in 79 research universities is calculated by means of a DEA model featuring inputs. From column 2 of Table 4, E_0 it can be seen that the efficiency value of decision units 2, 3, 4, 7, etc. is 1, and the lowest efficiency value is for decision unit 59, which is only 0.3778. Since the model (Equation 3) does not take into account the constraints of non-shared fixation and inputs, this paper further validates the calculations by using the method of 3. 2.

To determine the equilibrium efficiency frontier compliant with the non-shared fixed and variable input constraints as per synergy theory, Model (Equation 7) is utilized to compute the optimal adjustments for both the research and development system and the system for translating and applying research outcomes, with the results detailed in Table 3. A positive value in the table signifies that a region must augment its government grant input to align with the communal equilibrium efficiency frontier, while a negative value suggests a reduction is necessary, and zero indicates no adjustment is required. Table 3 reveals that 53 decision-making units require a decrease in government grant inputs, whereas 26 units necessitate an increase. The optimal adjustments, when summed, equal zero, fulfilling the criterion of maintaining a fixed total for government allocation inputs across all decision-making units [45].

The public equilibrium efficient frontier can be set based on the most favorable alteration and the efficiency of the two systems can be evaluated accordingly, and the results are shown in column 4 of Table 4. E_0^* shown. Compared with model (Equation 3), the efficiency results considering non-shared fixation and input constraints have changed dramatically, and the two-system R&D and R&D application efficiencies of the vast majority of decision units have been improved due to the non-shared fixation and constraints brought about by the government appropriation as an input to the R&D subsystem only, which makes the efficient frontier affected by the optimal adjustments of each decision unit, and no longer the efficient frontier as in the traditional method. The effective frontier surface in the traditional approach. In addition, the larger the optimal adjustment, the lower the efficiency of the two systems in terms of R&D and application of results. Model (Equation 9) also has the great advantage that the range of efficiency changes is much larger, which makes the R&D and R&D application efficiencies of each decision unit more distinguishable.

In terms of the amount of funding allocated, "985 Project" colleges and universities receive more funding than "non-985 Project" colleges and universities. For example, "985 Project" colleges include Decision Unit 1 (2,245), Decision Unit 3 (3140.1), Decision Unit 54 (2,561.9), Decision Unit 7 (1740.2), and Decision Unit 4 (3185.5). "Non-985 Project" universities such as Decision Unit 19 (144.19), Decision Unit 21 (207.91) and Decision Unit 55 (201.24). This shows that there is a significant gap in government funding between "985-project" and "non-985-project" colleges and universities.

Several insights can be gleaned from the data presented in Table 4. Firstly, there is a pronounced disparity in the efficiency of Chinese research universities. Certain decision-making units (DMUs), such as 4, 8, and 13, achieve the maximum efficiency score of 1, indicating optimal performance. In contrast, DMU 2 has a significantly lower efficiency score of only 0.2746. The average efficiency across the 79 universities surveyed is 0.7125. The presence of super-efficiency scores inflates this average, suggesting that the actual potential for improvement in real-world scenarios is likely to be greater than the 28.75% implied by the average efficiency score. Secondly, the performance of "Project 985" universities does not always meet expectations. For instance, DMU 2 has an efficiency score of 0.2746, DMU 1 scores 0.4690, and DMU 11 scores 0.4778. Typically, "985 Project" universities are known for their academic prowess and prestige and are often tasked with managing national key R&D programs. These programs tend to focus on critical science and technology issues related to industrial competitiveness, independent innovation capabilities, and national security. The lengthy and challenging nature of these research endeavors could explain the lower efficiency scores observed. Finally, there is no clear correlation between geographic location and efficiency. Certain institutions located in regions with less robust economic development have been observed to surpass the performance of similar universities in economically affluent areas. This suggests that the level of economic development in a region does not serve as a direct indicator of the progress of its research universities. However,

TABLE 2 Descriptive statistics of input-output indicators.

variant	FS	GS	GGF	RF	AS	PM	PP	TNS	TTI
Average value	3363	1785	797820	943482	25	21	5,021	229	19154
Upper quartile	2,229	1,185	501787	657965	18	13	3744	160	2,650
(Statistics) standard deviation	3089.03	1,660.94	789596.27	860347.20	22.44	18.43	3867.77	190.95	47196.96
Maximum values	13905	8,355	3185507	4638324	112	81	20701	1,012	302898
Minimum value	270	183	75925	69085	2	0	110	36	0

TABLE 3 Optimal adjustments.

DMU	Adjustment	DMU	Adjustment	DMU	Adjustment	DMU	Adjustment
1	−2,245.04	21	823.93	41	−295.38	61	595.43
2	862.70	22	−489.96	42	−296.55	62	511.60
3	−1,443.11	23	657.13	43	−1,007.55	63	820.71
4	542.88	24	−82.79	44	261.10	64	−965.09
5	439.15	25	468.82	45	441.08	65	−1988.59
6	−466.15	26	709.77	46	523.83	66	404.81
7	0.00	27	−799.32	47	570.61	67	83.21
8	219.19	28	−849.64	48	−8.02	68	780.68
9	663.74	29	772.71	49	701.24	69	320.07
10	−684.41	30	−326.40	50	657.61	70	−757.84
11	272.98	31	181.79	51	−2,436.83	71	−2,529.03
12	550.29	32	727.72	52	527.62	72	187.01
13	634.66	33	752.66	53	808.45	73	−512.10
14	647.53	34	−2,720.46	54	−1766.68	74	497.17
15	675.51	35	−981.53	55	596.24	75	337.79
16	392.95	36	−2,655.34	56	−85.20	76	520.03
17	−380.37	37	486.40	57	379.49	77	−552.67
18	760.53	38	684.12	58	367.29	78	476.65
19	712.41	39	404.03	59	−709.70	79	290.87
20	126.45	40	457.74	60	203.12		
Sum of Adjustment				0			

TABLE 4 Efficiency and ranking of model (Equation 3) and model (Equation 9).

DMU	E_0	Rank	E_0^*	Rank	DMU	E_0	Rank	E_0^*	Rank
1	0.5592	59	0.4690	70	41	0.7844	27	0.9660	18
2	1.0000	2	0.2746	76	42	0.4338	78	0.6381	51
3	1.0000	2	0.7783	32	43	0.4907	74	0.5582	61
4	1.0000	2	1.0000	1	44	0.7493	29	0.9336	21
5	0.6444	44	0.4390	73	45	0.6842	36	0.4447	72
6	0.8517	21	0.8644	27	46	0.6367	45	0.7094	38
7	1.0000	2	0.0000	79	47	0.6204	47	1.0000	1
8	1.0000	2	1.0000	1	48	0.5960	54	0.7078	39
9	0.6612	42	0.5634	60	49	0.6838	37	0.9094	23
10	0.7338	31	0.9635	19	50	0.6477	43	0.5028	65
11	0.6006	53	0.4778	68	51	0.6689	41	0.9710	17
12	0.5518	61	0.8519	28	52	0.5413	64	1.0000	1
13	0.9907	14	1.0000	1	53	0.9737	15	0.7932	31
14	1.0000	2	1.0000	1	54	0.5271	67	0.6656	47
15	0.7175	33	0.5936	57	55	0.8672	20	1.0000	1
16	0.5273	66	0.6350	52	56	0.6183	48	1.0000	1
17	0.6360	46	0.6159	55	57	0.6098	51	0.6695	46
18	0.9730	16	0.0000	79	58	0.5473	62	0.8660	26
19	0.8132	23	0.7571	34	59	0.3778	79	0.6472	49
20	0.4672	77	0.5508	62	60	0.7358	30	1.0000	1
21	0.7967	24	0.6730	44	61	0.7254	32	0.7094	37
22	0.5024	72	0.7337	35	62	1.1604	1	1.0000	1
23	0.5279	65	0.6218	54	63	0.8255	22	0.7180	36
24	0.8895	19	1.0000	1	64	1.0000	2	1.0000	1
25	1.0000	2	1.0000	1	65	0.5861	56	0.4338	74
26	1.0000	2	0.0000	79	66	0.5460	63	0.7047	40
27	0.5043	71	0.8143	29	67	0.6690	40	1.0000	1
28	0.6729	39	0.9197	22	68	0.7865	26	0.7755	33
29	0.9100	18	0.9355	20	69	0.5177	69	0.6842	43
30	0.6836	38	0.6081	56	70	1.0000	2	1.0000	1
31	0.9509	17	0.8931	24	71	0.5875	55	0.5121	64
32	1.0000	2	1.0000	1	72	0.4720	76	0.4736	69

(Continued on the following page)

TABLE 4 (*Continued*) Efficiency and ranking of model (Equation 3) and model (Equation 9).

DMU	E_0	Rank	E_0^*	Rank	DMU	E_0	Rank	E_0^*	Rank
33	0.9997	13	0.8871	25	73	0.7044	35	0.5812	58
34	0.5268	68	0.4930	67	74	0.7062	34	0.5420	63
35	0.5135	70	0.6727	45	75	0.5795	57	0.7039	41
36	0.6152	49	0.4680	71	76	0.4732	75	0.6445	50
37	0.6014	52	0.6323	53	77	0.6111	50	0.4229	75
38	0.7877	25	0.7990	30	78	0.7623	28	0.6856	42
39	0.5586	60	0.4974	66	79	0.5611	58	0.5750	59
40	0.4942	73	0.6538	48					

it is acknowledged that high-performing research universities can contribute to the economic development of their local regions.

5 Conclusion

This paper develops a specialized DEA method, grounded in synergy theory, to evaluate the efficiency of Chinese research universities, with a focus on outcome transformation. The model surmounts previous limitations by accounting for non-shared fixed and variable inputs within a network structure, providing a comprehensive assessment of DMUs in complex systems.

The study recommends that universities adhere to collaborative innovation theory to enhance research efficiency, involving all stakeholders. Governments should support and safeguard scientific achievements, while enterprises should leverage innovation to increase the application value of research. Financial institutions should back scientific endeavors, and technology transfer organizations should facilitate resource integration and innovation. A collaborative system involving these entities can improve evaluation services and drive economic progress through scientific applications.

To optimize Government Grant Funds (GGF) allocation, central administrators must accurately assess research activities and implement efficient funding mechanisms. Despite the suboptimal efficiency of some "985 Project" universities, their pursuit of advanced science and technology involves significant investments and risks. Administrators should monitor tech trends and understand the long cycles of innovation.

In conclusion, this study presents an innovative DEA approach for assessing Chinese research universities' performance in outcome transformation, acknowledging the need for future research to address technological diversity and radial measure limitations in DEA methods.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

YM: Writing–review and editing. MX: Writing–original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research is supported by the Double First-class Discipline Construction Project (PA2024GDGP0032) in Hefei University of Technology, the Provincial Quality Engineering Project of Colleges and Universities (2023jyjxggyjY034) and the Hefei University of Technology Teaching Quality and Teaching Reform Project (JYQN2216).

Acknowledgments

The authors are also grateful to Professor Tao Ding's guidance on this paper.

Conflict of interest

Author YM was employed by Anhui Sanxiang Technology Consulting Co., Ltd.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Lehmann EE, Meoli M, Paleari S, Stockinger SA. Approaching effects of the economic crisis on university efficiency: a comparative study of Germany and Italy. *Eurasian Business Rev* (2018) 8:37–54. doi:10.1007/s40821-017-0091-7

2. Martínez-Campillo A, Fernández-Santos Y. The impact of the economic crisis on the (in) efficiency of public Higher Education institutions in Southern Europe: the case of Spanish universities. *Socio-Economic Plann Sci* (2020) 71:100771. doi:10.1016/j.seps.2019.100771

3. Chen J. *Collaborative innovation*. Hangzhou, Zhejiang, China: Zhejiang University Press (2012).

4. Cook WD, Liang L, Zhu J. Measuring performance of two-stage network structures by DEA: a review and future perspective. *Omega* (2010) 38(6):423–30. doi:10.1016/j.omega.2009.12.001

5. Kao C, Hung HT. Efficiency analysis of university departments: an empirical study. *Omega* (2008) 36(4):653–64. doi:10.1016/j.omega.2006.02.003

6. Kao C, Lin PH. Qualitative factors in data envelopment analysis: a fuzzy number approach. *Eur J Oper Res* (2011) 211(3):586–93. doi:10.1016/j.ejor.2010.12.004

7. Azizi H, Jahed R. Improved data envelopment analysis models for evaluating interval efficiencies of decision-making units. *Comput and Ind Eng* (2011) 61(3):897–901. doi:10.1016/j.cie.2011.04.017

8. Villano RA, Tran CDT. Performance of private higher education institutions in Vietnam: evidence using DEA-based bootstrap directional distance approach with quasi-fixed inputs. *Appl Econ* (2018) 50(55):5966–78. doi:10.1080/00036846.2018.1488077

9. Lu WM. Intellectual capital and university performance in Taiwan. *Econ Model* (2012) 29(4):1081–9. doi:10.1016/j.econmod.2012.03.021

10. Chang TY, Chung PH, Hsu SS. Two-stage performance model for evaluating the managerial efficiency of higher education: application by the Taiwanese tourism and leisure department. *J Hospitality, Leis Sport and Tourism Education* (2012) 11(2):168–77. doi:10.1016/j.jhlste.2012.04.003

11. An Q, Yang M, Chu J, Wu J, Zhu Q. Efficiency evaluation of an interactive system by data envelopment analysis approach. *Comput and Ind Eng* (2017) 103:17–25. doi:10.1016/j.cie.2016.10.010

12. Zhang L, Guo C, Wei F. Multistage network data envelopment analysis: semidefinite programming approach. *J Oper Res Soc* (2019) 70(8):1284–95. doi:10.1080/01605682.2018.1489348

13. Yang GL, Fukuyama H, Song YY. Measuring the inefficiency of Chinese research universities based on a two-stage network DEA model. *J Informetrics* (2018) 12(1):10–30. doi:10.1016/j.joi.2017.11.002

14. Wang X, Hu H, Xie C. Auditing the efficiency of the nation-funded social science research with data envelopment analysis. *INFOR: Inf Syst Oper Res* (2019) 57(2):165–86. doi:10.1080/03155986.2018.1533208

15. Ding T, Yang J, Wu H, Wen Y, Tan C, Liang L. Research performance evaluation of Chinese university: a non-homogeneous network DEA approach. *J Management Sci Eng* (2021) 6(4):467–81. doi:10.1016/j.jmse.2020.10.003

16. Chen Y, Ma X, Yan P, Wang M. Operating efficiency in Chinese universities: an extended two-stage network DEA approach. *J Management Sci Eng* (2021) 6(4):482–98. doi:10.1016/j.jmse.2021.08.005

17. Xiong B, Wu J. Performance measurement in the parallel interdependent processes systems under decentralized and centralized modes. *J Oper Res Soc* (2021) 72(11):2442–59. doi:10.1080/01605682.2020.1796534

18. Xiong X, Yang GL, Zhou DQ, Wang ZL. How to allocate multi-period research resources? Centralized resource allocation for public universities in China using a parallel DEA-based approach. *Socio-Economic Plann Sci* (2022) 82:101317. doi:10.1016/j.seps.2022.101317

19. Chen X, Liu X, Zhu Q, Jiang J. R&D innovation efficiency of Chinese high-tech industries: three-stage network approach with fairness consideration. *J Oper Res Soc* (2022) 73(7):1562–77. doi:10.1080/01605682.2021.1920346

20. Moon H, Min D. Assessing energy efficiency and the related policy implications for energy-intensive firms in Korea: DEA approach. *Energy* (2017) 133:23–34. doi:10.1016/j.energy.2017.05.122

21. Alizadeh R, Beiragh RG, Soltanisehat L, Soltanzadeh E, Lund PD. Performance evaluation of complex electricity generation systems: a dynamic network-based data envelopment analysis approach. *Energ Econ* (2020) 91:104894. doi:10.1016/j.eneco.2020.104894

22. Lins MPE, Gomes EG, de Mello JCCS, de Mello AJRS. Olympic ranking based on a zero sum gains DEA model. *Eur J Oper Res* (2003) 148(2):312–22. doi:10.1016/S0377-2217(02)00687-2

23. Gomes EG, Lins MPE. Modelling undesirable outputs with zero sum gains data envelopment analysis models. *J Oper Res Soc* (2008) 59(5):616–23. doi:10.1057/palgrave.jors.2602384

24. Li F, Wang Y, Liang L. A two-stage efficiency evaluation method for DEA with shared output fixed and constrained. *China Management Sci (in Chinese)* (2023) 31(11):268–78. doi:10.16381/j.cnki.issn1003-207x.2021.1432

25. Ding T, Zhang Y, Zhang D, Li F. Performance evaluation of Chinese research universities: a parallel interactive network DEA approach with shared and fixed sum inputs. *Socio-Economic Plann Sci* (2023) 87:101582. doi:10.1016/j.seps.2023.101582

26. Altundemir ME. The impact of the financial crisis on American public universities. *Int J Business Soc Sci* (2012) 3(8):190–8. doi:10.6017/ihe.2012.3.8.21

27. Marire J. Are South African public universities economically efficient? Reflection amidst higher education crisis. *South Afr J Higher Education* (2017) 31(3):116–37. doi:10.20853/31-3-1037

28. Wolinsky H. The crash reaches the universities: the global financial crisis threatens private and public university funding in the USA and Europe. *EMBO Rep* (2009) 10(3):209–11. doi:10.1038/embor.2009.17

29. Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Sci* (1984) 30(9):1078–92. doi:10.1287/mnsc.30.9.1078

30. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *Eur J Oper Res* (1978) 2(6):429–44. doi:10.1016/0377-2217(78)90138-8

31. Li F, Zhang D, Zhang J, Kou G. Measuring the energy production and utilization efficiency of Chinese thermal power industry with the fixed-sum carbon emission constraint. *Int J Prod Econ* (2022) 252:108571. doi:10.1016/j.ijpe.2022.108571

32. Chen L, Guo M, Li Y, Liang L, Salo A. Efficiency intervals, rank intervals and dominance relations of decision-making units with fixed-sum outputs. *Eur J Oper Res* (2021) 292(1):238–49. doi:10.1016/j.ejor.2020.10.024

33. Bian Y, Hu M, Xu H. Measuring efficiencies of parallel systems with shared inputs/outputs using data envelopment analysis. *Kybernetes* (2015) 44(3):336–52. doi:10.1108/k-04-2014-0067

34. Li L, Dai Q, Huang H, Wang S. Efficiency decomposition with shared inputs and outputs in two-stage DEA. *J Syst Sci Syst Eng* (2016) 25:23–38. doi:10.1007/s11518-016-5298-0

35. Liu JB, Zhang X, Cao J, Chen L. Mean first-passage time and robustness of complex cellular mobile communication network. *IEEE Trans Netw Sci Eng*(2024) 11:3066–76. doi:10.1109/tnse.2024.3358369

36. Liu JB, Zheng YQ, Lee CC. Statistical analysis of the regional air quality index of Yangtze River Delta based on complex network theory. *Appl Energ*(2024) 357:122529. doi:10.1016/j.apenergy.2023.122529

37. Liu JB, Wang X, Cao J. The coherence and properties analysis of balanced 2p-Ary tree networks. *IEEE Trans Netw Sci Eng*(2024) 11:4719–28. doi:10.1109/tnse.2024.3395710

38. Yang M, Li YJ, Liang L. A generalized equilibrium efficient frontier data envelopment analysis approach for evaluating DMUs with fixed-sum outputs. *Eur J Oper Res* (2015) 246(1):209–17. doi:10.1016/j.ejor.2015.04.023

39. Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. *Eur J Oper Res* (2002) 142(1):16–20. doi:10.1016/s0377-2217(01)00293-4

40. Yang F, Wu DD, Liang L, O'Neill L. Competition strategy and efficiency evaluation for decision making units with fixed-sum outputs. *Eur J Oper Res* (2011) 212(3):560–9. doi:10.1016/j.ejor.2011.02.029

41. Yang M, Li Y, Chen Y, Liang L. An equilibrium efficiency frontier data envelopment analysis approach for evaluating decision-making units with fixed-sum outputs. *Eur J Oper Res* (2014) 239(2):479–89. doi:10.1016/j.ejor.2014.05.013

42. Danowski JA. Identifying collaborative innovation networks: at the inter-departmental level. *Procedia-Social Behav Sci* (2010) 2(4):6404–17. doi:10.1016/j.sbspro.2010.04.050

43. Johnes J, Li YU. Measuring the research performance of Chinese higher education institutions using data envelopment analysis. *China Econ Rev* (2008) 19(4):679–96. doi:10.1016/j.chieco.2008.08.004

44. An Q, Wang Z, Emrouznejad A, Zhu Q, Chen X. Efficiency evaluation of parallel interdependent processes systems: an application to Chinese 985 Project universities. *Int J Prod Res*(2019) 57(17):5387–99. doi:10.1080/00207543.2018.1521531

45. Liu JB, Liu BR, Lee CC. Efficiency evaluation of China's transportation system considering carbon emissions: evidence from big data analytics methods. *Sci Total Environ* (2024) 922:171031. doi:10.1016/j.scitotenv.2024.171031

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 19 March 2025
DOI 10.3389/fphy.2025.1562015

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Zijian Feng,
Beijing Normal University, China
Xu Fu,
Sichuan University, China

*CORRESPONDENCE
Chao Wang,
wangchaoedu@njau.edu.cn

RECEIVED 16 January 2025
ACCEPTED 27 February 2025
PUBLISHED 19 March 2025

CITATION
Ma Y, Sun Q and Wang C (2025) Systemically important company identification based on topological characteristics of the common share pledge network.
Front. Phys. 13:1562015.
doi: 10.3389/fphy.2025.1562015

COPYRIGHT
© 2025 Ma, Sun and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Systemically important company identification based on topological characteristics of the common share pledge network

Yuyin Ma[1], Qiong Sun[2] and Chao Wang[1]*

[1]School of Economics and Management, Southeast University, Nanjing, China, [2]School of Finance, Nanjing Agricultural University, Nanjing, China

Introduction: With the continuous development of the global securities market, share pledges as a business form are gradually emerging. However, there are also certain risks associated with share pledges. There are many common pledges between different share pledgers, which may result in financial contagion.

Methods: This paper constructs a common share pledge network and proposes an identification method of systemically important listed companies by analyzing the characteristics of its network structure, to improve the stability of the financial market.

Results: The results show that the share pledges are concentrated in a certain number of large companies, exhibiting scale-free characteristics and multiple local community features. The supervision of share pledges will promote the formation of such characteristics. Although financial supervision reduces the probability of share pledge risks, it also improves the importance of companies with high centrality in the share pledge market.

Discussion: Compared with the traditional network characteristics the centrality of the common share pledge network will provide more information about the share pledge risk. Different characteristics of network centrality can effectively identify systemically important companies, which has important regulatory value for financial risk prevention.

KEYWORDS

share pledge, network structure, systemic importance, stock market, financial risk

1 Introduction

Share pledge refers to an act of shareholders using their own equity holdings as collateral to finance funds from financial institutions such as securities companies or banks [1]. Share pledges not only provide an important way for shareholders to raise funds but are also seen as a loan-like business model offered by banks and securities companies, which has resulted in significant earnings growth for financial institutions [2]. In recent years, the business model of share pledging in emerging markets has seen unprecedented growth. In developed capital markets, share pledges are rare due to the relatively decentralized distribution of equity. In contrast, listed companies in emerging markets tend to exhibit a more concentrated shareholding structure, which leads to greater financing challenges for companies. Therefore, share pledges have become more prevalent in emerging markets. Due to the risk factors

associated with this financing method, the possibility of contagion risk has also increased.

Empirical evidence most share pledges hurt stock prices [3]. When a company discloses a share pledge, it usually conveys negative information to the market about its poor operation and financing difficulties, which may lead to a drop in the stock price. To mitigate these negative effects, the probability of capitalization following a share pledge by a major shareholder also increases [4]. Although this may help stabilize the stock price in the short term, it does not really improve the long-term business performance of the company and is even likely to give rise to speculative behaviors such as market value management [5, 6]. With high pledge ratios, large shareholders will be more inclined to pursue high-risk investment projects to earn more returns [7], which not only pose a financial risk but also involve moral hazard. Due to information asymmetry, shareholders also have incentives to manipulate stock prices. This moral hazard may lead to a crash in stock prices. If the stock price falls sharply, the pledgee may ask major shareholders to buy back the equity in advance, thus increasing the risk of control power transfer [8–10].

The share pledge business is a common yet high-risk financing approach. To effectively manage the risk of default due to stock price volatility, the pledger usually sets multiple risk control mechanisms such as pledge rate, warning line, and liquidation threshold. If the pledging ratio of major shareholders on the exchange market is higher, the mark-to-market and mandatory liquidation mechanisms are more likely to be activated during stock price declines. It will exacerbate the leverage effect and lead to a much higher stock price crash risk [1]. It is relatively common for pledged shares in emerging markets to hit liquidation thresholds, which has attracted extensive regulatory attention [11]. When the stock market encounters continuous downturn pressure, if the pledger fails to fill the account, or the stock price falls below the preset liquidation thresholds, banks and other financial institutions may take measures to mitigate the potential loss by "forced liquidation" for the sake of risk control and asset preservation [12]. Due to the common share pledge network, a crisis at an institution has the potential to spread to other institutions. If an institution suffers forced liquidation, its equity value will decrease, thereby affecting other institutions holding shares in that institution. To minimize losses, these institutions may choose to sell the shares at low prices [13]. According to signaling theory, such selling behavior may trigger a chain reaction in the market, further exacerbating the decline in corporate performance and elevating credit risk and default risk. The more connected institutions are in the share pledge network, the faster this risk contagion occurs [1].

For this risk contagion process, complex networks and network dynamics provide a useful tool for the study [14–19]. Tong et al. (2021) constructed a network model of major shareholders of listed companies and found that the degree of common shared shareholders accelerates the risk in the stock market [20]. In addition, the impact of common shareholders consisting of financial institutions is more pronounced. Greenwood et al. (2015) explored how systemic risk is triggered by selling investments based on the common investment network in the face of negative shocks [21]. These studies not only explore the risk contagion mechanism in financial markets but also provide assistance in the formulation of risk management and regulatory policies.

In summary, it is important to study the impact of contagion risk on companies due to share pledges. Network relationships in share pledges provide an effective way to study contagion risk between companies. Therefore, this paper portrays the common network relationship between listed companies with share pledges to study the risk contagion process. It further identifies the systemic importance of listed companies based on topological characteristics of the common share pledge network. This study not only helps to reveal the risk of share pledges but also identifies systemically important companies. In addition, this study provides a new perspective for regulatory policies and intervention strategies to maintain the stability of the capital market.

2 Methodology

2.1 Share pledge network

Share pledge defaults involve a large number of stakeholders with intricate relationships with each other. The existence of a common share pledge relationship can contribute to a stock price collapse when shares pledges are sold by the pledgee. The basis for portraying common share pledge relationships among listed companies is a bipartite network between listed companies and institutional investors.

The bipartite network contains two types of nodes and connections between nodes. It can be expressed as $G = \{Q, F, V\}$, where Q is the set of listed companies (pledgers), F is the set of institutional investors (pledgees), and V denotes the connections of the listed companies and the institutional investors through the relationship of share pledges. Corresponding to the bipartite network, its adjacency matrix can be expressed as $V = \left\{V_{ij}\right\}_{N\times M}$, where N and M denote the number of pledgees and pledgers in the share pledge network. When a financial institution holds a pledged share in a company, a network connection is established between the two types of nodes. The connection between the two types of nodes represents the share pledge relationship between the listed company and the financial institution. If the pledgee j is related to the pledger i, then $V_{ij} = 1$, otherwise $V_{ij} = 0$.

2.2 Common share pledge network

Based on the share pledge bipartite network, the traditional network model is further used to portray a common share pledge network between different companies [22]. The common share pledge network is constructed as follows: taking all the companies pledging shares as nodes, if there exists at least one common pledgee in any company, they are defined to have a common share pledge relationship, i.e., there exists a network connecting edge. The common share pledge network can be represented by $C = \{Q, L\}$, where $Q = \{1, 2, \cdots, N\}$ denotes the number of listed companies. $L = \{L_{ik}\}_{N\times N}$ is the adjacency matrix of the network, where L_{ik} represents whether node i and node k are connected. If two companies i and j have a common share pledge relationship, then $L_{ik} = 1$, otherwise $L_{ik} = 0$.

2.3 Topological characteristics of the network

The main topological characteristics of networks include the connectivity of the nodes and the distribution of connections. The connectivity of the common share pledge network contains network density, average degree, clustering coefficient, average shortest path, and modularity.

Network density reflects the tightness of connections between nodes in a network. It is defined as the ratio of the number of actually existing edges to the number of all possible edges. The network density is calculated by

$$d = \frac{2L}{N(N-1)} \tag{1}$$

where L is the number of edges actually present in the network.

The average path length is the average of all of the distances over all pairs of nodes in a network, which reflects the overall connectivity of the network. It also measures the efficiency with which information travels through the network because the efficiency in a network is also defined by distances [23, 24]. The average path length is calculated by

$$h_i = \frac{1}{N(N-1)} \sum_{k \neq i}^{N} h_{ik} \tag{2}$$

The clustering coefficient measures the degree of interconnection between neighboring nodes in a network, reflecting local connectivity and the "clique" phenomenon. The clustering coefficient is one of the local characteristics for networks that are based at the node level, it can be calculated by

$$c_i = \frac{2e_i}{k_i(k_i-1)} \tag{3}$$

Modularity is a measure of quality that a network is partitioned into modules (or communities). It reflects the extent to which nodes in the network are classified into modules with tight internal connections but sparse connections to other modules. Modularity is calculated by

$$q_i = \frac{1_i}{2L} \sum_{k \neq i}^{N} \left(L_{ik} - \frac{k_i k_k}{2L} \right) I_{\{m_i, m_k\}} \tag{4}$$

where m_i and m_k denote the modules to which node i and node k belong, and I is an indicative function that is 1 when node i and node k belong to the same module and 0 otherwise. The algorithm and resolution of modularity are referred to Blondel et al. (2008) and Lambiotte et al. (2008) respectively [25, 26].

In addition, network centrality is often used to identify the important nodes in the network, which is a basis for identifying systemically important companies in common share pledge networks [27]. Network centrality is classified as degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality. Degree centrality and network density portray similar network characteristics. The same as closeness centrality and average path length. Therefore, only other centrality metrics are considered.

Betweenness centrality measures the degree to which a node mediates the connection relationship between any two other nodes in the network. It is defined as

$$z_i = \sum_{k,s \in Q}^{N} \frac{\theta_{ks}(i)}{\theta_{ks}} \tag{5}$$

where θ_{ks} is the total number of all shortest paths from node k to node s. $\theta_{ks}(i)$ represents the total number of all shortest paths from node k to node s after the network intervenes at node i.

Eigenvector centrality evaluates the centrality of a node in the network based on the importance of its neighbors. It mainly considers the importance of vertices as passing bridges in the graph, rather than a centrality measure that only considers the direct importance of vertices. It is defined as

$$t_i = \frac{1}{\lambda} \sum_{k \in Q}^{N} L_{ik} t_k \tag{6}$$

where $t = (t_1, t_2, \cdots, t_N)$ is the eigenvector of the adjacency matrix L and λ is the corresponding eigenvalue.

2.4 Risk characteristics of companies

The systemic risk of the share pledge market is portrayed by the dynamic network analysis. Based on the common share pledge network, the risk contagion process is simulated by the system dynamics method. Assume that the company i with a common share pledge relationship suffers a crisis. Its market value is defined as 0 to simplify the simulation. Then the financial institution s holding the share pledge of the company i will liquidate these assets. Moreover, the financial institution s may also sell its other pledged shares in exchange for liquidity due to risk aversion. Given that the company j has a share pledge relationship with the financial institution s, then it will also suffer contagion risks due to the stock price drops caused by the selling of pledged shares. Based on the market depth of the stock, we define that the proportion of sold share pledges to the total market value of the company j is equal to the proportion of its stock price decline. Therefore, if the pledged shares of the company i are sold by all of its pledgees, the stock price of the company j may experience a notable decline. Moreover, if the company j has a share pledge relationship with other financial institutions at the same time, the contagion risk will further spread to other financial institutions. However, financial institutions will set a liquidation threshold for the pledged shares to prevent the stock price from falling and causing losses. We define that the liquidation threshold is 130%, which is the common standard in the Chinese stock market. It means the proportion of the share pledge losses when approaching the liquidation threshold is $1 - 1.3r$, where r is the share pledge ratio of a company. Therefore, when the stock price of the company j has declined to the liquidation threshold, its pledgees will sell the pledged shares in the stock market. This will trigger a domino effect until the share pledge market reaches equilibrium. The cumulative market value loss of all sold pledged shares is defined as the dynamic contagion risk. Based on the dynamic risk contagion process, 1,000 experiments were simulated to ensure that each company suffered the crisis in theory. The systemic risk is measured by the ratio of the dynamic contagion risk to the total size of the share pledge market, and the average value is used as the final result.

The individual risks of listed companies are portrayed by the default distance. The default probability model is used to estimate the individual default risk of listed companies [28]. It is calculated as

$$dd_{it} = \frac{log\left(\frac{E_{it}+D_{it}}{D_{it}}\right) + \left(r_{it-1} - \frac{\sigma_{Vit}^2}{2}\right)T_{it}}{\sigma_{Vit}\sqrt{T}_{it}} \quad (7)$$

where dd_{it} denotes the default distance of company i at moment t; E_{it} denotes the total market capitalization of the company that is calculated by the product of the total number of shares issued and the market price; D_{it} denotes the face value of the company's debt that obtained by adding half of the company's long-term liabilities to its short-term liabilities; $r_{it\text{-}1}$ denotes the company's annual log stock return; σ_{Vit} denotes the company's asset volatility that calculated from the volatility of the company's stock price return σ_{it}, that is

$$\sigma_{Vit} = \frac{E_{it}\sigma_{it}}{E_{it}+D_{it}} + \frac{D_{it}}{E_{it}+D_{it}}(0.05 + 0.25\sigma_{it}) \quad (8)$$

We include the five percentage points in this term to represent term structure volatility, and we include the 25% times equity volatility to allow for volatility associated with default risk according to Bharath and Shumway (2008) [28]. Combining Equations 7, 8 we can calculate the company's default distance dd_{it}. On this basis, the company's default probability $p_{it} = N(-dd_{it})$ is calculated through the standard cumulative normal distribution function $N(\cdot)$.

3 Results and discussion

3.1 Data description

China is one of the emerging markets that require mandatory disclosure of stock pledges which provides appropriate data on share pledges [29]. Therefore, this study takes the annual data of share pledges of all listed companies in China's A-shares as a sample. The sample interval is chosen as 2008 to 2024. The sample includes a total of 3,312 listed companies, all of which have engaged in new share pledge business within the sample interval. According to the statistics of share pledge data, the average share pledge term is about 1.39 years, so the time window is set to 1 year. Data on share pledges are obtained from the CSMAR database. To highlight the role of financial institutions in the risk contagion process, it constructs the common share pledging network by excluding the relationship in which an individual investor serves as the pledgee and the network relationship formed with institutional investors such as banks and securities companies as pledgees are considered.

The common share pledge network is uniquely determined by the node and the edge. Therefore, the characterization of nodes and edges is the key to revealing the relationship of the common share pledge network. This study mainly explores the following systemic characteristics of the common share pledge network: density, average degree, clustering coefficient, average shortest path, and modularity, which are calculated by Equations 1-6 respectively. These economic implications are mainly divided into two categories, which describe the node characteristics of the network (e.g., clustering coefficient, modularity) and edge characteristics (e.g., average degree, average shortest path). The network characteristics were calculated by Gephi 0.10.1. When calculating these characteristics, the unconnected nodes were excluded because they have no contribution to the contagion process of the share pledge risk.

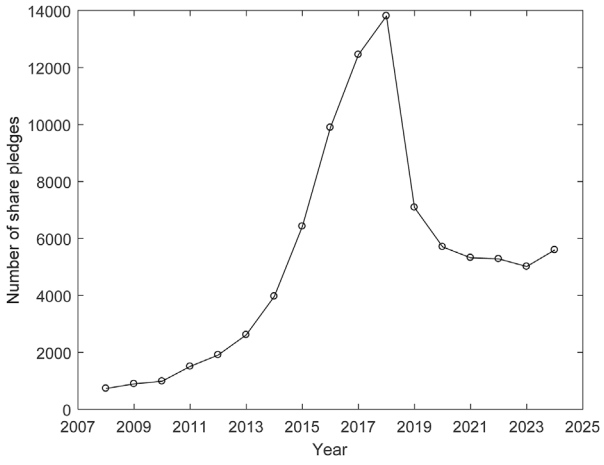


FIGURE 1
Number of share pledges.

Figure 1 shows the statistical results of the number of share pledges of all listed companies in China's A-shares from 2008 to 2024. The share pledges in China's stock market have gradually become prevalent since 2008, especially after the opening of exchange-based share pledges in 2013. Because the market had no unified standards for the limit of the pledge rate, the number of share pledges increased dramatically. After the share pledge crisis broke out in 2018, the financial regulatory authorities issued policies to restrict pledgees from forcibly liquidating share pledges. Meanwhile, with the regulator's requirement in 2018 that the cap on the pledge ratio of listed companies' shares should not exceed 60%, the financing way of share pledges began to contract sharply. The number of share pledges has gradually and steadily declined. Due to the impact of COVID-19 on the global economy and society from 2019, the number of share pledges has continued to shrink and it is still at a low level at present.

3.2 Systemic topological characteristics

The analysis of topological characteristics of the common share pledge network helps to explore the trends of China's share pledge market. The number of connected edges of the network is the most direct characteristic, which is mainly portrayed by the density and degree distribution of the network. Network density in Figure 2 mainly measures the degree of connectivity of the entire network. Notably, the network density and degree distribution are calculated without considering isolated nodes. The isolated node lacks a channel for contagion risks, so it is not helpful for risk identification. In terms of common share pledge relationships, Figure 2 shows apparent phased characteristics. The network density was relatively stable from 2008 to 2012. After the opening of exchange-based share pledges in 2013, the network density showed a substantial increase and remained at a high level from 2013 to 2017. It implies that as

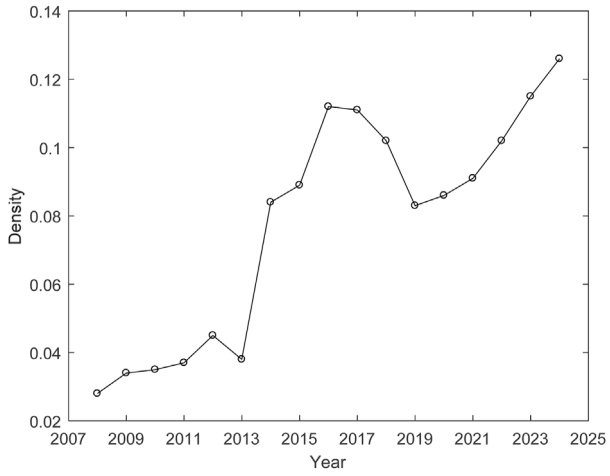


FIGURE 2
Density of the common share pledge network.

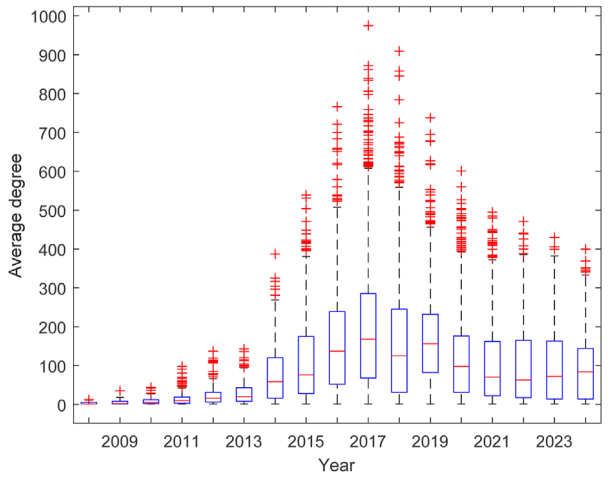


FIGURE 3
Average degree of the common share pledge network.

the number of share pledges increases, the pledgees are gradually concentrating on some large financial institutions (as shown in Section 3.3). The phased density has shown a substantial decline as a result of strengthening regulations following the outbreak of the share pledge crisis in 2018. However, unlike the downward trend of the number of share pledges in Figure 1, the density of the common share pledge network continues to rise and exceed the historical peak after COVID-19 in 2019. Due to the high-risk property of share pledges, most financial institutions gave up share pledges as an investment strategy after the elimination of COVID-19. Therefore, the current share pledge is mainly concentrated in some large financial institutions. This is an inevitable trend in market development, but it also increases the concentration risk of share pledges, which provides a warning for future regulation.

The average degree of the network in Figure 3 mainly portrays the average number of participants in common share pledges per listed company. Due to the limited number of pledgees, more pledgers will inevitably choose the same pledgees when share pledges become more prevalent. It will lead to an increase in the number of common share pledges in the market. Figure 3 shows that the listed company in China's stock markets has the highest common share pledge relationship with nearly 200 other listed companies. It means that the common share pledge in China's stock market is widely prevalence. In addition, Figure 3 shows a trend consistent with Figure 1 i.e., common share pledges were prevalent with the regulation liberalization in 2013–2017 and seriously controlled after the tightening of financial regulations in 2018. Although the average degree is stably around 100, it is still at a high level, so the supervision of the common share pledge still needs to be strengthened.

We further analyze the characteristics of the common share pledge network in terms of average path length, clustering coefficients, and community structure. All of them portray the degree of local associations in the network. The results are shown in Figures 4–6. The trends in average path length, clustering coefficients, and modularity are generally similar which are opposite to the trends in the connectivity of the network. Overall, the average

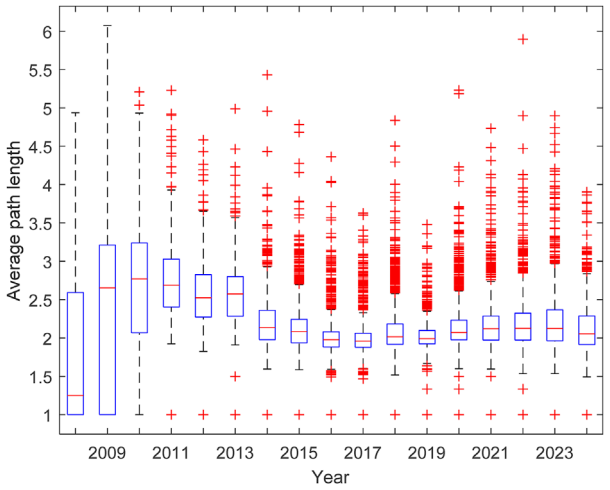


FIGURE 4
Average path length of the common share pledge network.

path lengths, clustering coefficients, and modularity of the network show a downward trend over the 2013–2017 period, while they all increase after the strengthening of financial regulation since 2018. However, they have fallen back after the COVID-19.

The average path length portrays the accessibility between nodes. The longer the average path the more distant the network relationship between two nodes. As the number of share pledges increases, the likelihood that companies have a common share pledge relationship with each other is promoted, which greatly reduces the distance between different companies. Therefore, the average path length shows an opposite trend to the average degree of the common share pledge network. The clustering coefficient portrays the degree of agglomeration between different nodes. Although the average degree of the common share pledge network increased between 2013 and 2017, the trend of the clustering coefficient decreased in Figure 5. However, the average degree of

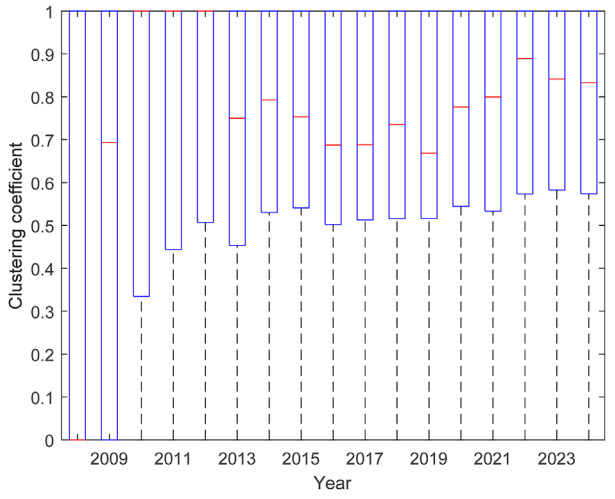


FIGURE 5
Clustering coefficient for the common share pledge network.

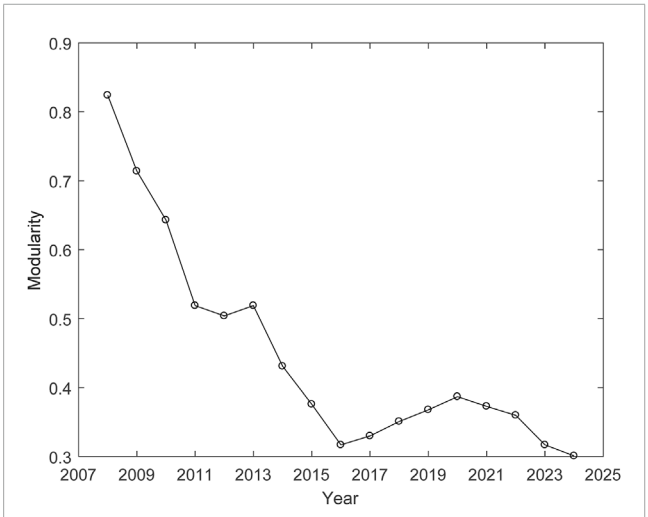


FIGURE 6
Modularity of the common share pledge network.

the common share pledge network decreased from 2018 to 2024, while its clustering coefficient increased instead. This indicates that participants are widely active during the booming period of the share pledge market, while only some large financial institutions are active in the face of financial recession. The characteristics of the modularity of the common share pledge network in Figure 6 further validate this conclusion. Modularity mainly characterizes the quality of community structure division in a network, with higher modularity indicating stronger community structure. If a company in the community suffers a crisis, its share pledges held by the financial institution will be liquidated. In this case, other companies in the community will also be impacted. Therefore, the modularity provides channels for risk contagion and helps to identify key nodes in the community. In summary, there are some differences in the topological characteristics among different nodes in the common

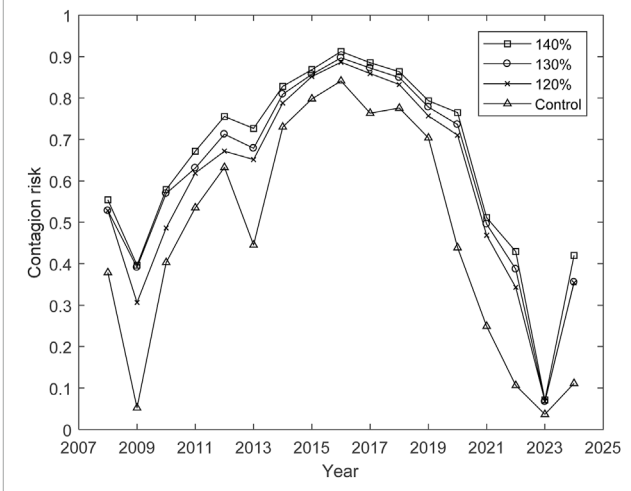


FIGURE 7
Contagion risk of the common share pledge network.

share pledge network. These differences provide risk information for identifying systemically important companies.

The topological structure of the common share pledge network helps to mine the risk information of listed companies, and then identify systemically important companies. The predictive capacity of network characteristics is further validated by the dynamic contagion risk of the share pledge market in Figure 7. Figure 7 presents that the potential contagion risk gradually increased with the increasing number of share pledges from 2008 to 2012. Especially after the liberalization of on-exchange share pledges in 2013, the contagion risk in the share pledge market continued to rise and remained at a high level. Due to the supervision of the share pledge market in 2018 and the following shock of COVID-19, the number of share pledges dropped dramatically. As a result, the contagion risk of share pledges has been controlled. However, with the end of COVID-19, the contagion risk has rebounded in recent years. The dynamic contagion risk is closely related to the density of the common share pledge network, indicating that the common share pledge network provides a channel for the contagion risk. Therefore, the topological characteristics implied by the common share pledge network are helpful for the risk identification of listed companies. Moreover, the highest proportion of contagion risk accounts for about 80%–90% of the share pledges, indicating that not all nodes will suffer contagion risks. This is mainly because some nodes with low modularity or high average paths are almost independent, resulting in an extremely long channel of risk contagion.

Figure 7 also presents the simulation results of contagion risk at the 120% and 140% liquidation thresholds to enhance the robustness. They show the same trend as the baseline case (130% liquidation threshold), and the level of contagion risks are very close. It means the results of contagion risks are robust. The robustness analysis in Figure 7 also indicates that a higher liquidation threshold is more likely to trigger fire sales of the share pledges. Therefore, the Chinese stock market has set a 20% higher warning line before the liquidation threshold to indicate the share pledge risk and reduce concentrated fire sales. In addition, this study further explores the impact of reducing the share pledge ratio on contagion

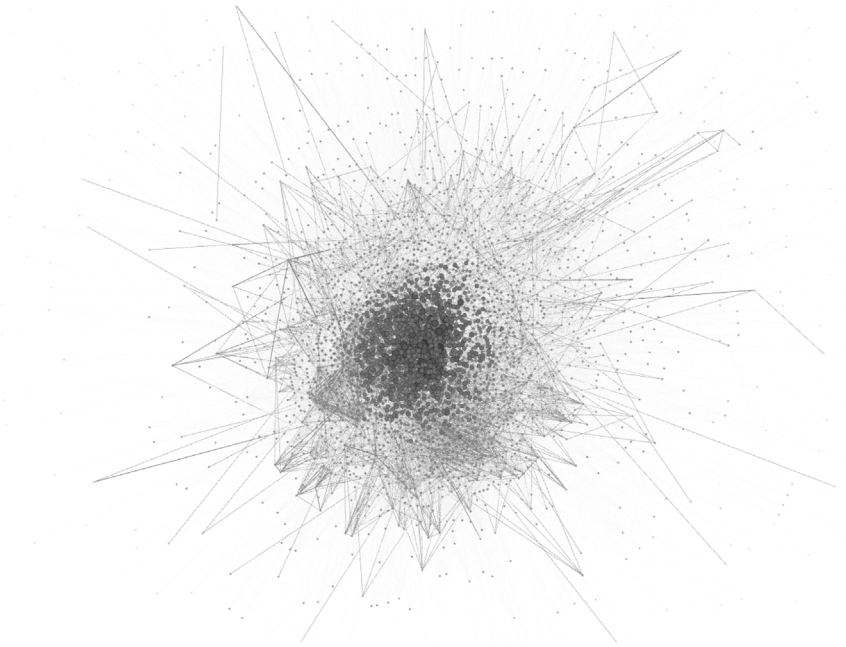


FIGURE 8
Common share pledge network relationship.

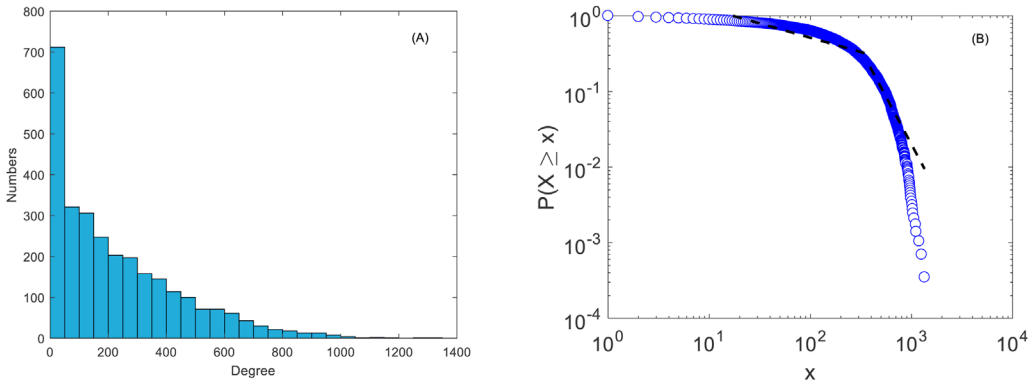


FIGURE 9
Individual degree distribution of the common share pledge network. **(A)** Histogram plots; **(B)** Logarithm plots.

risks to discuss the risk control strategies. The result is labeled as "Control" in Figure 7, which assumes that the share pledge ratio is 80% of the original level. A similar result indicates that the model is effective. Moreover, reducing the share pledge ratio is more effective in controlling contagion risks when compared to the results of different liquidation thresholds, which highlights that strengthening supervision contributes to financial stability.

3.3 Individual topological characteristics

Further analyzing the individual topological characteristics of the common share pledge network, which are portrayed in terms of degree, clustering coefficient, and average path length. In addition, eigenvector centrality and authority centrality are also the metrics to portray the individual network structure.

To capture the stock of unresolved share pledges, this study analyzes the individual topological characteristics of the common share pledge network using accumulated share pledge data from 2008 to 2024. Figure 8 visualizes the network, where the nodes represent the listed companies in China's stock markets, and their colors indicate the number of common share pledge relationships that exist in the company. The connections indicate the existence of a common share pledge relationship between companies, and their thickness indicates the number of common pledgees between any two companies. Figure 8 shows that most of the listed companies have few common pledgees in their share pledge relationships, and many companies have dense common pledgees,

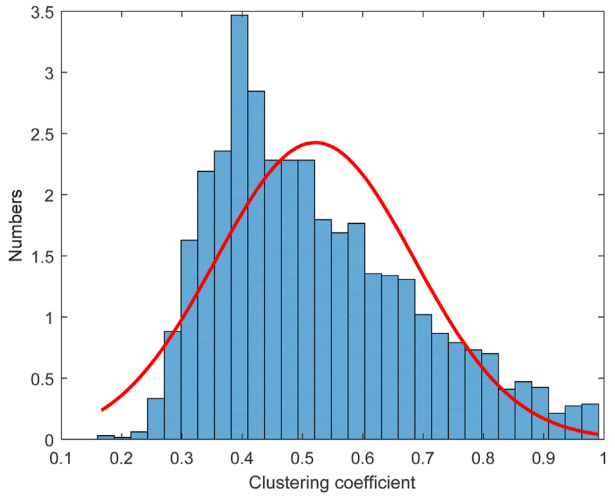


FIGURE 10
Individual clustering coefficient for the common share pledge network.

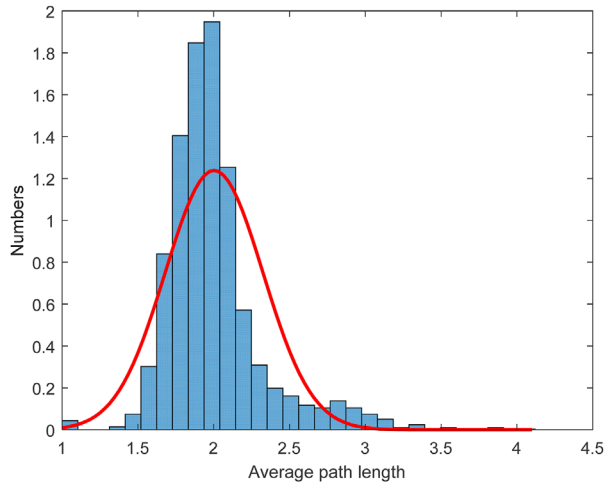


FIGURE 11
Individual average path length for the common share pledge network.

thus forming a network characterized by multiple "core-periphery" distributions. This conclusion is also emphasized by the individual degree distribution with "leptokurtosis and heavy tails" presented in Figure 9A. Figure 9B further presents the individual degree distribution in double logarithmic axes. Although the individual degree distribution of the network is difficult to fit with a single function, it can be expressed as a power law distribution with an exponential truncation, i.e., it is characterized by a scale-free distribution.

The individual statistics of clustering coefficients and average path lengths are shown in Figures 10, 11. They show that the clustering coefficients and average path lengths of the nodes are generally fitted by normal distributions. The clustering coefficient fits a normal distribution with a mean of 0.5211 and a standard deviation of 0.1644. However, the distribution of clustering coefficients does not exactly conform to the standard normal distribution but rather has the characteristics of a positively skewed distribution. Recall that the clustering coefficients of the common share pledge network in Figure 5 are high, while that formed by the undischarged share pledges in Figure 10 is lower. It means that the cycles of most share pledges are short and the common share pledge relationship is highly concentrated in the short term, thus forming the phased feature of the share pledge relationship.

The average path length is fitted with a normal distribution with a mean of 2.0025 and a standard deviation of 0.3221. The average path length is not exactly in line with the standard normal distribution. It also fits the lognormal distribution with "leptokurtosis and heavy tails," which indicates that the common share pledges of some listed companies are denser.

The statistical results of centralities are shown in Figures 12, 13. Although different types of centrality distributions have differences, they are similar to moderate centrality in Figure 9 and both show scale-free distributions. It further indicates that some listed companies have dense common share pledges. Therefore, systemically important nodes can be identified based on the density of common share pledges.

3.4 Regression results

It further mines the risk information of listed companies from the individual topological characteristics of the network and conducts regression analysis for the default probability of individual companies. In addition to the classical clustering coefficient and average shortest path in complex networks, the individual characteristics related to the centrality (i.e., the degree centrality, betweenness centrality, and eigenvector centrality) are considered to identify systemically important listed companies. To eliminate the impact of other explicit factors, control variables are set as follows. The fundamental variables of the company include logarithmic total assets (size) and share concentration (hhi). The financial performance variables include a debt-to-asset ratio (lev) and return on assets (roa). The market performance variables include annual turnover rate (hls) and price-to-book ratio (pb). The companies labeled by ST were excluded from the sample to eliminate the influence of outliers. The missing data were removed from the regression model. In addition, the winsorization is used at the 1% level to test the robustness of regressions.

Table 1 presents a variable correlation matrix to evaluate potential multicollinearity issues of the variables. The result indicates that there is a low correlation between the dependent variables, explanatory variables, and control variables. However, the correlation between centrality variables is high, which is consistent with the fact. Then the Variance Inflation Factor (VIF) values are further calculated. It shows an average VIF value of 1.85, and the VIF values of each variable are all less than 2. Therefore, it is no potential multicollinearity issues among the variables in this model. Table 2 presents statistical descriptions of all variables. It shows that there are differences in the statistical characteristics of different centrality variables, and this is the basis for characterizing the network topology through different centralities.

This study analyzes the impact of individual topological characteristics of the network on the default probability of

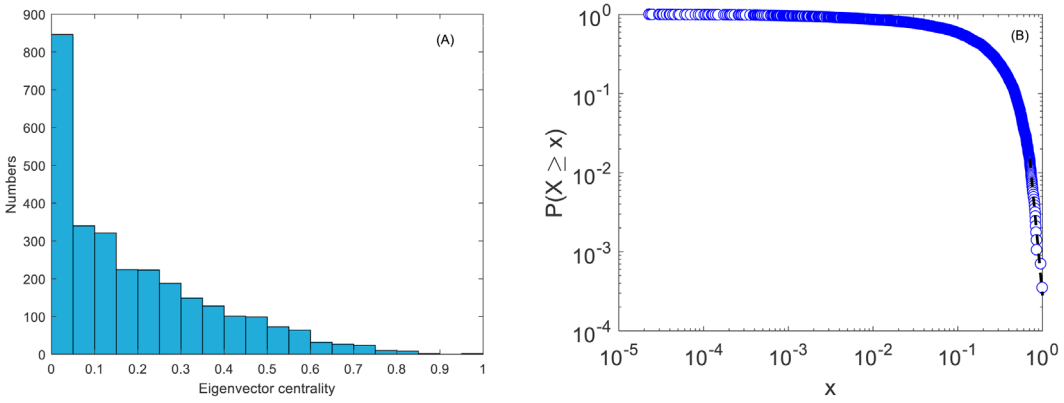


FIGURE 12
Eigenvector centrality of the common share pledge network. **(A)** Histogram plots; **(B)** Logarithm plots.

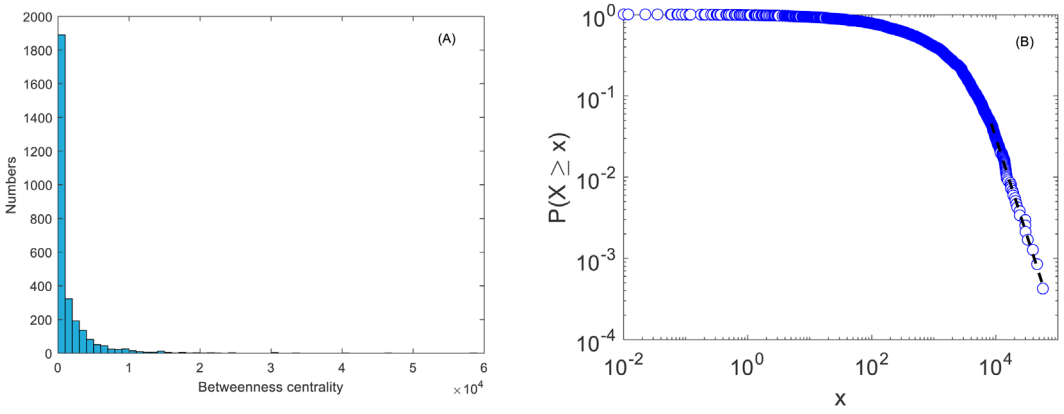


FIGURE 13
Betweenness centrality of the common share pledge network. **(A)** Histogram plots; **(B)** Logarithm plots.

companies using the two-way fixed effects model in the baseline regression. The results of Table 3 help to explore the effectiveness of the common share pledge network in identifying systemically important nodes. The regression coefficients of degree centrality, betweenness centrality, and eigenvector centrality with companies' default probability are all positively significant at the 1% level, and that of the average path length is negatively significant at the 10% level. Given the fact that the dense common share pledge relationships promote the risk exposure between the pledgor and pledgee, thereby it increases the contagion risk faced by the companies. This indicates that the risk information of companies can be effectively identified through the centrality characteristics of the network. The critical nodes or critical paths identified based on the centrality characteristics provide channels for contagion risk among listed companies and increase the systemic risk that companies face. However, the regression coefficient of the clustering coefficient is negatively significant at the 1% level, which is consistent with the result of Figure 5. It indicates that companies with higher clustering coefficients do not mean higher contagion risks, while centrality characteristics are more able to reflect the contagion risk of the share pledge market compared to traditional network properties. We further use a 1% winsorized sample for regression to test the robustness of the model, and the robust results are presented in Table 4.

As the risk of stocks increases, the space for shareholders to "profit" from the stocks they hold is continuously compressed and gradually disappears, and even greater wealth losses may occur. Therefore, when the stock risk of a company increases, shareholders are more likely to pledge their equity, converting equity with "market value" into highly liquid funds with "usable value" as soon as possible, thereby transferring the risk of wealth loss due to stock price decline and liquidity risk to the pledgees. This process highlights the endogeneity issue of our study. In addition, the calculation of default probability requires the use of stock volatility. Therefore, the lag of the explained variable should be considered in terms of the previous period's stock volatility. To further address the endogeneity issue of our study, the system Gaussian mixture model (GMM) is further used for regression, and the results are presented in Table 5. The regression coefficients of GMM generally verify the explanatory of centrality characteristics

TABLE 1 Variable correlations.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
dp	1.000										
degree	0.007	1.000									
betweeness	0.018	0.669	1.000								
clustering	0.026	−0.269	−0.399	1.000							
path length	−0.015	−0.521	−0.278	0.134	1.000						
eigenvector	0.024	0.832	0.556	−0.185	−0.477	1.000					
size	−0.302	0.097	0.189	−0.132	−0.062	0.081	1.000				
lev	0.151	0.028	0.013	0.021	0.003	0.050	0.058	1.000			
roa	−0.507	−0.032	0.032	−0.067	0.009	−0.017	0.215	−0.559	1.000		
pb	0.096	−0.013	−0.004	0.006	0.000	−0.010	−0.127	−0.086	0.026	1.000	
hsl	0.048	−0.041	−0.104	0.115	−0.016	−0.004	−0.364	0.025	−0.121	0.073	1.000
hhi	0.002	−0.064	0.005	−0.017	0.046	−0.027	0.055	0.081	0.009	−0.017	−0.001

TABLE 2 Statistical descriptions.

Variables	N	Mean	Std	Min	Max
dp	10,280	0.011	0.078	0.000	0.700
degree	10,280	134.332	130.591	1.000	535.000
betweeness	10,280	889.304	1707.332	0.000	27,410.955
clustering	10,280	0.693	0.272	0.000	1.000
path length	10,280	2.138	0.435	1.000	6.037
eigenvector	10,280	0.232	0.224	0.000	1.000
size	10,280	22.184	1.164	18.900	26.848
lev	10,280	0.424	0.190	0.0600	0.876
roa	10,280	0.034	0.069	−0.332	0.200
pb	10,280	3.639	2.951	0.000	18.393
hsl	10,280	6.327	4.639	0.823	30.113
hhi	10,280	0.136	0.094	0.018	0.494

on the share pledge risk. The result highlights the robust conclusion after considering endogeneity issues.

According to the analysis, financial regulation has a decisive impact on the share pledge market. To explore how changes in the network density are related to policy adjustments, the policy dummy variable is used for the regression to provide relevant causal analyses. The policy dummy variable is conducted with 2018 as the threshold. The results after adding policy items and their interaction items in the baseline regression are presented in Table 6. The results indicate that the coefficients of regulatory policy are negatively significant with the default risk at the 1% level, suggesting that strengthening supervision of share pledges can effectively mitigate financial risks. The coefficients of interaction terms between regulatory policy and network centrality characteristics are significant at the 5% or 10% level. Moreover, they have the same symbol as network centrality characteristics, indicating that the role of network centrality characteristics in risk identification is more prominent when regulatory policies are strengthened. At this time, the share pledge market is shrinking, and some companies are eliminated by the market, which strengthens the network centrality characteristics of the remaining companies and their position in risk contagion. The policy interaction term of the clustering coefficient is not significant because its ability to identify risks is weak as discussed. This also emphasizes the importance of network centrality characteristics in risk identification.

We further incorporate a real-world case to enhance the practical relevance and policy implications of the study. At the end of 2018, the total share pledges of A-share reached ¥ 4.61 trillion, accounting for 10.32% of the total A-share market value. There are 3,542 companies in A-share with unlisted share pledges, accounting for 99.13% of the total number of A-share listed companies. Among them, over 60 companies have pledged more than 60% of their shares, exceeding the regulatory threshold of 50%.

This study takes Yinyi Co., Ltd. as the case to discuss the 2018 tightening of share pledge financing regulations in China and its impact on contagion risk. Yinyi Co., Ltd. (000981) is a real estate company listed on A-share. It was forced liquidation when faced with stock crashes during the 2018 market turmoil due to high

10.3389/fphy.2025.1562015

TABLE 3 Baseline regression for individual topological characteristics and default probability.

	dp	dp	dp	dp	dp
Degree centrality	0.0046***(0.0008)				
Betweenness centrality		0.0004***(0.0001)			
Eigenvector centrality			1.4420***(0.4493)		
Clustering coefficient				−0.6091*(0.3338)	
Average path length					−0.5662*(0.2204)
size	−1.2180***(0.2029)	−1.2395***(0.2028)	−1.1289***(0.2014)	−1.0959***(0.1997)	−1.1044***(0.1997)
lev	6.8980***(1.3697)	7.0412***(1.3692)	7.0198***(1.3705)	6.9986***(1.3709)	7.0700***(1.3703)
roa	−22.4735***(0.8639)	−24.4146***(0.8640)	−24.5015***(0.8647)	−24.5635***(0.8641)	−24.5110***(0.8645)
pb	0.5561***(0.0385)	0.5532***(0.0385)	0.5592***(0.0385)	0.5626***(0.0384)	0.5595***(0.0385)
hsl	−0.0622***(0.0240)	−0.0621***(0.0240)	−0.0652***(0.0240)	−0.0658***(0.0240)	−0.0667***(0.0240)
hhi	−0.0307 (1.9083)	−0.0956 (1.9077)	−0.2207 (1.9101)	−0.1482 (1.9097)	−0.1615 (1.9095)
Time fixed	Yes	Yes	Yes	Yes	Yes
Company fixed	Yes	Yes	Yes	Yes	Yes
N	10,280	10,280	10,280	10,280	10,280
R^2	0.3971	0.4014	0.3990	0.4004	0.3998

TABLE 4 Winsorized result for individual topological characteristics and default probability.

	dp	dp	dp	dp	dp
Degree centrality	0.0030***(0.0008)				
Betweenness centrality		0.0003***(0.0001)			
Eigenvector centrality			0.8989**(0.4108)		
Clustering coefficient				−0.5694*(0.3009)	
Average path length					−0.4900**(0.2116)
Controls	Yes	Yes	Yes	Yes	Yes
Time fixed	Yes	Yes	Yes	Yes	Yes
Company fixed	Yes	Yes	Yes	Yes	Yes
N	10,280	10,280	10,280	10,280	10,280
R^2	0.5387	0.5404	0.5409	0.5421	0.5415

share pledge ratios. The share pledges of Yinyi Co., Ltd. have been increasing since 2015, with the highest proportion reaching 86.32%. At the end of 2018, its proportion of share pledges was 72.63%, which is far higher than the threshold of 50% in the new regulations. The market value of Yinyi Co., Ltd.'s share pledges ranked the top three A-share companies. Therefore, Yinyi Co., Ltd. is a typical case for study.

With the outbreak of share pledge risks in 2018, the stock price of Yinyi Co., Ltd. has continuously hit the limit down, and most of the share pledges approach the warning line. Therefore, the Yinyi

10.3389/fphy.2025.1562015

TABLE 5 GMM regression results for individual topological characteristics and default probability.

	dp	dp	dp	dp	dp
Degree centrality	0.0055***(0.0017)				
Betweenness centrality		0.0007**(0.0003)			
Eigenvector centrality			0.0882 (1.5921)		
Clustering coefficient				−4.3870**(1.9889)	
Average path length					−0.6011*(0.3306)
L.dd	0.1508***(0.0239)	0.1638***(0.0245)	0.1529***(0.0236)	0.1403***(0.0229)	0.1461***(0.0229)
Controls	Yes	Yes	Yes	Yes	Yes
AR (1)	−12.06 (0.000)	−11.90 (0.000)	−12.07 (0.000)	−12.08 (0.000)	−12.07 (0.000)
AR (2)	−3.21 (0.001)	−3.23 (0.001)	−3.32 (0.001)	−3.48 (0.001)	−3.37 (0.001)
Sargan test	1801.74 (0.000)	1739.91 (0.000)	1819.15 (0.000)	1887.42 (0.000)	1898.78 (0.000)
Hansen test	580.89 (0.000)	557.36 (0.000)	567.76 (0.000)	639.25 (0.000)	669.39 (0.000)
Wald test	14,784.42 (0.000)	14,821.55 (0.000)	14,437.87 (0.000)	14,121.96 (0.000)	14,215.94 (0.000)
N	9,313	9,313	9,313	9,313	9,313

TABLE 6 Impact of regulatory policy on individual topological characteristics and default probability.

	dp	dp	dp	dp	dp
Degree centrality	0.0033***(0.0011)				
Betweenness centrality		0.0002***(0.0001)			
Eigenvector centrality			0.8886*(0.4775)		
Clustering coefficient				−0.2757 (0.4207)	
Average path length					−0.2392 (0.2743)
	−5.8744***(1.1056)	−5.7238***(1.1033)	−6.0546***(1.1217)	−5.3151***(1.1828)	−3.8095***(1.4550)
	0.0026**(0.0013)	0.0002**(0.0001)	1.1474*(0.6523)	−0.7995 (0.6138)	−0.8659**(0.4326)
Controls	Yes	Yes	Yes	Yes	Yes
Time fixed	Yes	Yes	Yes	Yes	Yes
Company fixed	Yes	Yes	Yes	Yes	Yes
N	10,280	10,280	10,280	10,280	10,280
R^2	0.3977	0.4000	0.3990	0.4006	0.3999

Co., Ltd. tried to avoid the liquidation risk by suspending trading. However, with the strengthening of financial regulation, the strictest suspension and resumption rules have been implemented. After the Yinyi Co., Ltd. was forced to resume trading, its stock price continued to hit the limit down. The sharp decline in the market value of share pledges has increased the liquidation risk of Yinyi Co., Ltd. In addition, the difficulty of financing has also increased after tightening the share pledge supervision. Yinyi Co., Ltd. had to default in the face of the cash crunch, and it further received default disposal notices from five pledgees. The default triggers

market panic. This concern has led to a fire sale by investors, which exacerbates the decline in stock prices. Thus, Yinyi Co., Ltd. went downhill and it finally filed for bankruptcy in 2019. Moreover, the risk also spreads to other listed companies in the YinYi Group, such as Guangxi Hechi Chemical Co., Ltd. and Ningbo Kangqiang Electronics Co., Ltd. Eventually, 16 companies under the Yinyi Group merged and restructured together in 2020 and renamed the company as "SENSTEED HI-TECH GROUP" in 2023. However, this reform did not improve the company's losses. Since 2018, SENSTEED HI-TECH GROUP has suffered consecutive losses of up to ¥ 15 billion.

The case of Yinyi Co., Ltd. shows that excessive reliance on share pledge financing carries serious potential risks. Although the strong regulation of share pledge financing has an impact on the stock market in the short term, it is conducive to market stability and healthy development by cleaning up the stock market in the long run. Therefore, it should continue to improve the regulatory rules to promote high-quality development of the stock market in the future.

4 Conclusion

With the expansion of the securities market, the share pledge as a business form experienced booming growth in emerging markets. However, the risk of share pledges should also warrant attention. The share pledge relationships are more concentrated in emerging markets. As a result, there are a large number of common share pledges between different share pledgers. Due to this common share pledge network, once the risk breaks out, it will have a cascading effect. The impact will spread to multiple markets such as the stock market, banks, securities, and funds, which in turn will bring great harm to the entire financial market and the real economy. Therefore, this study proposes an identification method for systemically important listed companies by analyzing topological characteristics of a common share pledge network to improve the stability of the financial system and the ability to regulate financial risks. The following conclusions are obtained from the study.

Share pledge relationships in emerging financial markets are characterized by phases. Most share pledges have short cycles, and the common share pledge relationships form multiple localized communities in the short term. The strengthening financial regulation will promote the concentration of share pledges in some large companies, and the common share pledge relationship is characterized by a scale-free distribution. Although it reduces the possibility of risks in the share pledge market, it also enhances the position of companies with high centrality characteristics in the contagion of share pledge risks. This phenomenon implies direction for future financial regulation. The centrality characteristics of the common share pledge network can provide more risk information for identifying systemically important companies, which is helpful in financial risk prevention.

The findings implicate the following policy insights. The regulators should increase regulation for common share pledge issues to identify and address potential risks. The risks of the common share pledge should be diversified through policy guidance. It should optimize the structure of the share pledge market and promote its healthy development. More stable share pledge relationships are helpful to reduce frequent share pledges and share pledge releases in the short term, thereby reducing market volatility. It is important to strengthen the supervision of systemically important companies in the future. Using the common share pledge network to identify systemically important companies helps to prevent companies from triggering systemic risks.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YM: Conceptualization, Methodology, Software, Writing–original draft, Writing–review and editing. QS: Resources, Writing–original draft, Writing–review and editing. CW: Writing–original draft, Writing–review and editing, Supervision, Validation, Visualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by Ministry of Education of Humanities and Social Science (No. 21YJC790108).

Acknowledgments

The authors are grateful for the funding support in the writing process. They thank the editors for their hard work in the article review process and the reviewers for their professional opinions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Qin X, Wang Z. Share pledge financing network and systemic risks: evidence from China. *J Banking Finance* (2023) 152:106871. doi:10.1016/j.jbankfin.2023.106871

2. Li W, Zhou J, Yan Z, Zhang H. Controlling shareholder share pledging and firm cash dividends. *Emerging Markets Rev* (2020) 42:100671. doi:10.1016/j.ememar.2019.100671

3. Dou Y, Masulis RW, Zein J. Shareholder wealth consequences of insider pledging of company stock as collateral for personal loans. *Rev Financial Stud* (2019) 32(12):4810–54. doi:10.1093/rfs/hhz034

4. Dudley E. Social capital and entrepreneurial financing choice. *J Corporate Finance* (2021) 70:102068. doi:10.1016/j.jcorpfin.2021.102068

5. Chen CC, Chen CD, Lien D. Financial distress prediction model: the effects of corporate governance indicators. *J Forecast* (2020) 39(8):1238–52. doi:10.1002/for.2684

6. Sun X, Lei Y. Research on financial early warning of mining listed companies based on BP neural network model. *Resour Pol* (2021) 73:102223. doi:10.1016/j.resourpol.2021.102223

7. Anderson R, Puleo M. Insider share-pledging and equity risk. *J Financial Serv Res* (2020) 58(1):1–25. doi:10.1007/s10693-020-00332-x

8. Chan K, Chen HK, Hu SY, Liu YJ. Share pledges and margin call pressure. *J Corporate Finance* (2018) 52:96–117. doi:10.1016/j.jcorpfin.2018.08.003

9. Pang C, Wang Y. Stock pledge, risk of losing control and corporate innovation. *J Corporate Finance* (2020) 60:101534. doi:10.1016/j.jcorpfin.2019.101534

10. Hu J, Zhong L, Qiao J. Equity pledge and debt financing of listed companies. *Finance Res Lett* (2024) 59:104771. doi:10.1016/j.frl.2023.104771

11. Zhou M, Li K, Chen Z. Corporate governance quality and financial leverage: evidence from China. *Int Rev Financial Anal* (2021) 73:101652. doi:10.1016/j.irfa.2020.101652

12. Luo W, Guo X, Zhong S, Wang J. Environmental information disclosure quality, media attention and debt financing costs: evidence from Chinese heavy polluting listed companies. *J Clean Prod* (2019) 231:268–77. doi:10.1016/j.jclepro.2019.05.237

13. Pichler A, Poledna S, Thurner S. Systemic risk-efficient asset allocations: minimization of systemic risk as a network optimization problem. *J Financial Stab* (2021) 52:100809. doi:10.1016/j.jfs.2020.100809

14. An P, Zhou J, Li H, Sun B, Shi Y. The evolutionary similarity of the co-shareholder relationship network from institutional and non-institutional shareholder perspectives. *Physica A: Stat Mech Its Appl* (2018) 503:439–50. doi:10.1016/j.physa.2018.02.183

15. Fang T, Wang D, Lin Z, Wang X. Extreme risk measurement for the oil and China's sectors system—network-based approach and machine learning methods. *Front Phys* (2023) 11:1292418. doi:10.3389/fphy.2023.1292418

16. Ren C, Zhu Z, Zhou D. Multi-quantile systemic financial risk based on a monotone composite quantile regression neural network. *Front Phys* (2024) 12:1484589. doi:10.3389/fphy.2024.1484589

17. Jia K, Yin L. From governance contagion to risk contagion: research on systemic risk contagion from the perspective of governance externality. *Front Phys* (2024) 12:1476000. doi:10.3389/fphy.2024.1476000

18. Shao H, Wang D, Zhou B. Economic policy uncertainty on stock market risk contagion: a network-based approach. *Front Phys* (2024) 11:1094659. doi:10.3389/fphy.2023.1094659

19. Jia Y, Chen H, Liu J, Wang X, Guo R, Wang X. Exploring network dynamics in scientific innovation: collaboration, knowledge combination, and innovative performance. *Front Phys* (2024) 12:1492731. doi:10.3389/fphy.2024.1492731

20. Tong H, Jia Z, Zhang M, Qi J. Analysis of stock-shareholder associated network based on complex network. *J Math Finance* (2021) 11(01):107–22. doi:10.4236/jmf.2021.111005

21. Greenwood R, Landier A, Thesmar D. Vulnerable banks. *J Financial Econ* (2015) 115(3):471–85. doi:10.1016/j.jfineco.2014.11.006

22. Jackson MO. A typology of social capital and associated network measures. *Social choice welfare* (2020) 54(2):311–36. doi:10.1007/s00355-019-01189-3

23. Latora V, Marchiori M. Efficient behavior of small-world networks. *Phys Rev Lett* (2001) 87(19):198701. doi:10.1103/PhysRevLett.87.198701

24. Strang A, Haynes O, Cahill ND, Narayan DA. Generalized relationships between characteristic path length, efficiency, clustering coefficients, and density. *Social Netw Anal Mining* (2018) 8:14–6. doi:10.1007/s13278-018-0492-3

25. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech Theor Exp* (2008) 2008(10):P10008. doi:10.1088/1742-5468/2008/10/P10008

26. Lambiotte R, Delvenne JC, Barahona M. *Laplacian dynamics and multiscale modular structure in networks* (2008). arxiv preprint arxiv:0812.1770. doi:10.48550/arXiv.0812.1770

27. Kuzubaş TU, Ömercikoğlu I, Saltoğlu B. Network centrality measures and systemic risk: an application to the Turkish financial crisis. *Physica A: Stat Mech its Appl* (2014) 405:203–15. doi:10.1016/j.physa.2014.03.006

28. Bharath ST, Shumway T. Forecasting default with the Merton distance to default model. *Rev Financial Stud* (2008) 21(3):1339–69. doi:10.1093/rfs/hhn044

29. Zhu B, Xia X, Zheng X. One way out of the share pledging quagmire: evidence from mergers and acquisitions. *J Corporate Finance* (2021) 71:102120. doi:10.1016/j.jcorpfin.2021.102120

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 25 March 2025
DOI 10.3389/fphy.2025.1557932

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Siyao Liu,
Chinese Academy of Sciences (CAS), China
Siyu Huang,
Guangxi University, China

*CORRESPONDENCE
Qishen Chen,
chenqishen@126.com
Jingwen Mao,
jingwenmao@263.net

RECEIVED 09 January 2025
ACCEPTED 13 February 2025
PUBLISHED 25 March 2025

CITATION
Zhang Y, Chen Q, Wang K, Guan Q, Xing J, Long T, Zheng G, Li M, Shang C and Mao J (2025) Evolutionary characteristics and community detection of the international metal minerals trade.
Front. Phys. 13:1557932.
doi: 10.3389/fphy.2025.1557932

COPYRIGHT
© 2025 Zhang, Chen, Wang, Guan, Xing, Long, Zheng, Li and Shang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evolutionary characteristics and community detection of the international metal minerals trade

Yanfei Zhang[1], Qishen Chen[1]*, Kun Wang[1], Qing Guan[2], Jiayun Xing[1], Tao Long[1], Guodong Zheng[1], Mei Li[1], Chenghong Shang[3] and Jingwen Mao[4]*

[1]Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing, China, [2]School of Information Engineering, China University of Geosciences (Beijing), Beijing, China, [3]School of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing, China, [4]MNR Key Laboratory of Metallogeny and Mineral Assessment, Institute of Mineral resources, Chinese Academy of Geological Sciences, Beijing, China

Introduction: With the development of global industrialization, metal minerals have become a global focal point of great power competition. An in-depth investigation into the evolution of global mineral trade structures, alongside an analysis of the relationships between metal minerals trade, resource endowments, industrialization, regional dynamics, and geopolitical competition, is crucial for nations to formulate effective trade policies and enhance the stability of global mineral trade development.

Methods: This study explores the evolution trends of the global metal minerals trade structure from 1990 to 2022 based on complex network and further detect community structure using the Infomap algorithm.

Results and discussion: The results show that (1) There is a general upward trend in global metal minerals trade from 1990 to 2022, which can be segmented into phases of slow, rapid, and moderate growth. (2) The two major trading circles formed in the early stage with Japan and Europe as the center have changed, forming a dual pattern with China as the super demand center and Australia as the super supply center. (3) China, Australia, the U.S., Japan, Brazil, and the European Union play key roles in shaping the global trade network, with the structure of the global metal minerals trade network primarily being driven by demand centers. (4) According to the economic trends and the evolving resource demands, the global trade structure will translate from a concentrated bipolar model to a diversified network with multiple trade centers. The conclusion of this research helps specify international policies and maintain supply chain resilience.

KEYWORDS

metal minerals, international trade, complex network, map equation framework, community detection

1 Introduction

Metal minerals are the raw materials for industrial development. With the development of global industrialization and the transfer of industrial centers, the trade structure of mineral resources has also changed. In recent years, black swan events such as the epidemic of the century, Sino-U.S. trade friction, the Russia-Ukraine conflict, and the Palestine-Israel conflict have occurred frequently, disrupting the stability of trade networks and impacting the trade of mineral products [1, 2]. For example, Russia is one of the largest exporters of copper resources, accounting for 4% of the world's refined copper production. However, due to the ongoing fermentation of the Russia-Ukraine conflict, Russia's copper exports have declined significantly, resulting in a rise in LME copper prices, and a supply shortage of 87,000 tons in the global refined copper market in 2023. Meanwhile, the import of copper from Russia by EU countries decreased by 79% compared with 2022, to 62,000 tons, and China's import of copper from Russia increased by 14.8% compared with 2022 to 366,000 tons. As a result, exploring the evolution of the community structure of international metal minerals trade networks is essential for governments to adjust their development strategies. Previous studies on metal minerals trade networks usually focused on a single mineral, which cannot show the whole picture of the trade structure of mineral products. The complex network theory often be applied to carry out the supply chain network structure and risk of minerals such as copper [3, 4], lithium [5], nickel [6, 7], cobalt [8–11]and boron [12]. Additionally, models and methods such as multi-layer complex network construction, infectious disease spread models, and link prediction are widely used to characterize the evolution of global mineral resource supply risks and trade security [13–17]. Complex networks are typically composed of a few highly significant large nodes and numerous smaller nodes with less significance, often exhibiting community clustering phenomena. Community detection is a critical approach for investigating the internal structure and relationships within complex networks [18]. Three main algorithm types are widely employed in community detection for complex networks. The first includes hierarchical clustering-based algorithms, such as the Lawworth and Girvan-Newman algorithms [19, 20], which are suitable for communities with tree network structures. The second type is the greedy algorithm and Louvain algorithm based on maximum modularity-degree calculation [21–23], which are suitable for fast detection of undirected large complex network communities. The third type is an infomap algorithm based on random walks and data compression. The structure of the trading community of a single mineral species is often too simple, and only a few relevant studies have been carried out [24–27], using the second type of algorithm mentioned above, and the effect and explainability of the community division are not ideal. The infomap algorithm is a quality function for optimizing community recognition on directed and weighted networks, uses data compression for community detection, known as map equation [28, 29], which can be used for network visualization and dynamic evolution of community structure, and has developed rapidly in recent years. It is particularly effective for revealing hierarchical and overlapping communities in large-scale, weighted and directed networks [30–32]. This provides a robust framework for understanding the functional organization and resilience of the global metal minerals trade network. Therefore, this paper attempts to apply the infomap algorithm to the community detection research of mineral trade structure.

This paper selects the total trade values of 17 types of metal minerals products (trade code 2,601~2,617) [33] as the research object. It first constructs a complex network of global trade of metal minerals products and then utilizes the map equation framework to detect and divide the community structure of international mineral trade. In the analysis of network community structure and evolution process, this paper combined the role importance of a single node and the characteristics of inter-regional trade structure analysis, to reveal the evolution process of metal minerals products trade structure.

In order to accurately divide trading communities, we integrate map equations with complex network theory into a new analytical framework. We conducted community surveys over multiple periods to explore the evolution of the global trade structure in metals and minerals. This article has two main contributions. First, unlike the traditional analysis of a single mineral, this paper analyzes all metal minerals with publicly available trade data, and for the first time explores the global trade structure of metal minerals from a holistic perspective. Secondly, this paper constructs a community network model of global metal minerals trade by infomap algorithm, analyzes the process of trade center transfer, and reveals the evolution characteristics of global metal minerals trade.

The remainder of this paper is organized as follows. Section 2 introduces the data source, structural indices of the complex network, and the algorithm for the map equation framework. Section 3 explores the top traders, important regions, important communities, and their evolution over time. Section 4 discusses the findings, and Section 5 offers conclusions.

2 Data and methods

2.1 Data

All metal minerals trading data including the Harmonized System(HS) code from 2601 to 2617 (Table 1) were obtained from the UN Comtrade database (https://comtradeplus.un.org/), covering the period of 1990, 1995, 2000, 2005, 2010, and 2015-2022. This study has retrieved 72,311 records (Table 2), each of which includes exporters, importers, and trade value (the unit is dollars). These minerals collectively form the basis of our research object, which is to investigate the overall trade structure of metal minerals. Based on the data of 13 years, this study adds the total trade value of each metal minerals as the trade weight in each year, as shown in Equation 1:

$$W = \frac{1}{2}\sum_{i=1}^{n} w_{i,j} \tag{1}$$

where W represents the total trade value of all metal minerals in each year, $w_{i,j}$ represents the trade values of each metal mineral commodity between countries i and j . n is the total number of countries. m is the total number of metal mineral commodities.

Except the trading data, commodity price, GDP, GDP Per Capita, Manufacturing value added, minerals resource reserve, and production data are also used to analyze and interpret the results of the study, and the sources of these data are the World Bank (https://

TABLE 1 The HS code and the commodity name.

HS code	Commodity	HS code	Commodity	HS code	Commodity
2,601	Iron ores and concentrates	2,607	Lead ores and concentrates	2,613	Molybdenum ores and concentrates
2,602	Manganese ores and concentrates	2,608	Zinc ores and concentrates	2,614	Titanium ores and concentrates
2,603	Copper ores and concentrates	2,609	Tin ores and concentrates	2,615	Niobium, tantalum, vanadium or zirconium ores and concentrates
2,604	Nickel ores and concentrates	2010	Chromium ores and concentrates	2,616	Precious metal ores and concentrates
2,605	Cobalt ores and concentrates	2,611	Tungsten ores and concentrates	2,617	Ores and concentrates; n.e.c. in heading no. 2601
2,606	Aluminium ores and concentrates	2,612	Uranium or thorium ores and concentrates		

TABLE 2 Trading records of each year.

	1990	1995	2000	2005	2010	2015	2016	2017	2018	2019	2020	2021	2022	Total
Records	2,115	3,114	4,384	4,943	5,952	6,724	6,188	6,478	6,621	6,599	6,161	6,670	6,362	72,311

data.worldbank.org/), and United States Geological Survey (https://www.usgs.gov/centers/national-minerals-information-center).

2.2 Methodology

2.2.1 The construction of complex networks

Complex networks are ideal for analyzing trade patterns due to their ability to capture intricate interdependencies, non-linear interactions, and systemic structures in large-scale systems [34, 35]. This approach reveals key players, community dynamics, and resilience, offering insights beyond traditional models [36]. Compared with other methods, its strength lies in handling mutil-dimensional data and providing global perspectives, making it essential for understanding the evolution and efficiency of global trade.

Complex network model $G = (V, E)$ consists of node set (V) and edge set (E). In $V = \{1, 2, \ldots, n\}$, n represents the number of nodes, and in $E = \{1, 2, \ldots, m\}$, and m represents the number of edges. The complex network matrix is shown in Equation 2:

$$G = (V,E) = \begin{bmatrix} 0 & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & 0 & \ldots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \cdots & 0 \end{bmatrix} \quad (2)$$

where w_{ij} represents the edge weights of nodes i to j.

Therefore, we use the complex network model to construct a total of 13 directed weighted networks for the international trade of metal ores and concentrates from 1995 to 2022, in which the trading country is the node, the trade relationship is the edge, and the sum of all metal ores and concentrates trade volume is the edge weight.

2.2.2 Indices based on topological characteristics

Based on the complex network model of international metal minerals trade patterns, structural indicators are always used to measure the global or local characteristics from different perspectives. In this study, according to the topological features of the nodes and the global network, we selected five indies to measure the change in the community structure. Predicated on the difference in the pattern structure, these indices can reflect the impact of countries in the trade patterns when removing specific countries or regions. Simplified definitions and descriptions are as follows.

2.2.2.1 Degree

The degree represents the total number of edges connected between nodes in a network. In a directed network, the degree is composed of in-degree and out-degree representing the total export and import relations between countries. The calculations are shown in Equations 3–5:

$$k_i^{in} = \sum_{j=1}^{N} a_{ji} \quad (3)$$

$$k_i^{out} = \sum_{j=1}^{N} a_{ij} \quad (4)$$

$$K_i = k_i^{in} + k_i^{out} \quad (5)$$

where a_{ji} represents the edge from node j to node i and a_{ij} represents the edge from node i to node j.

2.2.2.2 Weighted degree

The weighted degree of nodes is the sum of the trade value of all the edges. In a directed network, the weighted degree can be divided into the weighted in-degree and the weighted out-degree, which are

shown by Equations 6–8:

$$s_i^{in} = \sum_{j=1}^{n} W_{ji} \tag{6}$$

$$s_i^{out} = \sum_{j=1}^{n} W_{ij} \tag{7}$$

$$S_i = s_i^{in} + s_i^{out} \tag{8}$$

where W_{ji} represents the edge weights of node j pointing to node i and W_{ij} represents the edge weights of node i pointing to node j.

2.2.2.3 Centrality

The centrality measures the importance and influence of nodes in a network and identifies the major active information channels. It is one of the core indices of network analysis. In this chapter, two types of centrality are considered: closeness centrality and betweenness centrality [37–39].

2.2.2.4 Betweenness centrality

The betweenness centrality measures the importance of a node to the shortest paths through the network, which describes the capability of a node that make connections to other pairs of nodes in a network. It is measured by Equation 9:

$$bc(i) = \sum_{j,k \in \nu V} \frac{\varsigma_{jk}(i)}{\varsigma_{jk}} \tag{9}$$

where $\varsigma_{jk}(i)$ is the number of shortest paths that pass through the node i, and ς_{jk} represents the total number of the shortest paths between any pair of nodes in the network.

2.2.2.5 Closeness centrality

The closeness centrality is the inverse of the average shortest distances from a node to all other nodes in a network. It measures how fast a node will take to spread information from the node to every other node that describe the influence of the node. It is defined as Equation 10:

$$c(i) = \frac{N-1}{\sum_{\forall c i \neq j} \rho(i,j)} \tag{10}$$

Where $\rho(i,j)$ represents the shortest path distance of node i and node j in the network.

In brief, weighted degree, weighted in-degree, and weighted out-degree are used to describe countries' total trade value, import, and export. The closeness centrality and betweenness centrality are selected to estimate the influence and importance of the network.

2.2.3 Communities detection with the map equation framework

The map equation framework is an information-theoretic method that identifies communities by optimizing the random walker's movement across the network to minimize its description length [29]. It is particularly effective for revealing hierarchical and overlapping communities in large-scale, weighted and directed networks [30–32]. This provides a robust framework for understanding the functional organization and resilience of the global metal minerals trade network. For a given network partition, denoted as M, the map equation specifies the theoretical limit $L(M)$ of how concisely we can describe the trajectory of this random walk.

To capitalize on the regional structure of the network, one index codebook and m module codebooks are utilized, with each module in the network having its codebook, to articulate the movements of the random walker. The module codebooks contain codewords for nodes within each module, including exit codes to depart the module, derived from the node visit/exit frequencies of the random walker. The index codebook comprises codewords for the modules, derived from the module switch rates of the random walker. Thus, the average length of the code describing a step of the random walker is the average length of the codeword from the index and module codebooks, weighted by their rates of use, as shown in Equation 11:

$$L(M) = q_{\curvearrowleft} H(\mathcal{Q}) + \sum_{i=1}^{m} p_{\circlearrowleft}^{i} H(\mathcal{P}^{i}) \tag{11}$$

The two-level average description length for a single step of the random walker within a network comprising n nodes, segmented into map M with m distinct modules, is composed of two components. The first term is the average description length of the index codebook, as shown in Equations 12, 13. The second term is the average description length of the module codebooks, as shown in Equations 14, 15.

$$q_{\curvearrowleft} = \sum_{i=1}^{m} q_{i\curvearrowleft} \tag{12}$$

The $q_{\curvearrowleft}$ describes the rate at which the index codebook is used. The per-step utilization rate of the index codebook is thus calculated as the sum of the probabilities with which the random walker accesses each module.

$$H(\mathcal{Q}) = -\sum_{i=1}^{m} \frac{q_{i\curvearrowleft}}{q_{\curvearrowleft}} \log \frac{q_{i\curvearrowleft}}{q_{\curvearrowleft}} \tag{13}$$

The $H(\mathcal{Q})$ describes the frequency-weighted average length of codewords in the index codebook. The entropy of the usage rates for the module codebooks quantifies the minimal average codeword length that can be theoretically achieved.

$$\sum_{i=1}^{m} p_{\circlearrowleft}^{i} = \sum_{i=1}^{m} \left(\sum_{\alpha \in i} p_{\alpha} + q_{\curvearrowright} \right) \tag{14}$$

The $\sum_{i=1}^{m} p_{\circlearrowleft}^{i}$ describes the rate at which the module codebooks are used. The per-step utilization rate of the module codebooks is dictated by the cumulative usage rate of the m module codebooks. For module i, this rate reflects the time spent by the random walker in the module, which is the sum of the probabilities of visiting any node in the module and using the exit message when leaving.

$$H(\mathcal{P}^{i}) = -\frac{q_{\curvearrowright}}{q_{\curvearrowright} + \sum_{\beta \in i} p_{\beta}} \log \frac{q_{\curvearrowright}}{q_{\curvearrowright} + \sum_{\beta \in i} p_{\beta}} - \sum_{\alpha \in i} \frac{p_{\alpha}}{q_{\curvearrowright} + \sum_{\beta \in i} p_{\beta}} \log \sum_{\beta \in i} p_{\beta} \tag{15}$$

The $H(\mathcal{P}^{i})$ describes the frequency-weighted average length of codewords in module codebook i. The entropy of the exit and visit rates of the random walker for module i represents the theoretical minimum of the average codeword length.

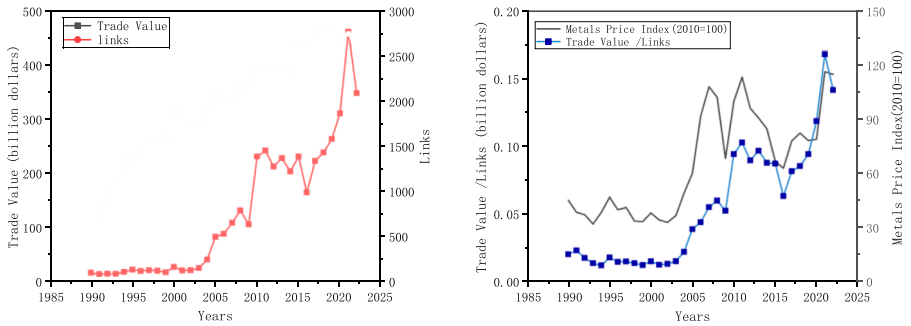


FIGURE 1
The trade value and metals price index, 1990–2022.

3 Results

3.1 Overview of global metal minerals trade

3.1.1 3 stages of global metal mineral trade

To investigate the evolutionary process of the global metal minerals trade structure, this study analyzes historical trade data from 1990 to 2022 in terms of trade volume, trade value, unit trade price, and the metal price index (Figure 1). According to the growth rate of value, and combined with the reality of global economy and mining development, this study divides the global metal minerals trade into three stages: (A) the slow development stage (1990–2000), (B) the rapid growth stage (2000–2010), and (C) the deceleration stage (2010–2022).

During Stage A (1990–2000), trade links doubled to 1949 in 2000, but trade value grew slowly, with new links contributing low value. Unit trade value declined slightly, while metal prices remained stable, indicating gradual trade volume expansion.

In Stage B (2000–2010), trade link growth slowed, but trade value surged nearly sevenfold to $232 billion, driven by rising unit trade values and metal prices, reflecting significant growth in both trade volume and value.

During Stage C (2010–2022), trade links stabilized (2,600–2,800), while trade value, unit value, and metal prices grew steadily. In 2021, the global economy rebounded post-pandemic, leading to a rise in metal prices and a new peak in trade value. By 2022, trade links declined and stabilized as industrial production normalized.

Overall, global trade in metals and minerals shows an upward trend in trade links, value, unit price, and the metal price index. Especially, the evolution process of trade value and unit price is similar.

3.1.2 Dominance countries evolution

Figure 2 illustrates the evolution of the global metal minerals trade network from 1990 to 2022, highlighting the weighted import and export of countries. It shows that a small percentage, about 10%–20%, of countries are responsible for 90% of the cumulative weighted indegree, while 11%–16% account for 90% of the cumulative weighted outdegree. This indicates that a minority of nations dominate global trade in metal minerals.

This paper examines the shifting dominance in global metal minerals trade, focusing on the top importers by the proportion of trade value annually (Figures 2C, D). From 1990 to 2000, Japan led trade until 2000, when its demand stabilized post-industrialization. Then China emerged as the top 1 importer, with its import growth mirroring its industrialization phase—rapid at first, then slowing as industrialization neared completion. Since 2010, the top 10 importers have consistently accounted for around 85% of global imports, indicating a high concentration of trade, with the proportion of top importer was near 90% by 2022. This suggests that a small number of countries dominate metal minerals imports.

Compared with the change of the number of top countries, the cumulative weighted export ratio is stable, mainly caused by the change of the top 5 major exporting countries, among which Australia, Brazil and Chile are stable major exporting countries. Analyzed the evolution of the cumulative weighted outdegree ratio among the top 30 countries since 1990, there is a gradual decline, indicating an increasing participation of countries in the export trade of metal minerals. Nonetheless, the preponderance of exports remains concentrated within the top 10 exporting nations.

3.1.3 Dominance commodities evolution

The global metal minerals trade from 1990 to 2022 was primarily driven by iron and copper, constituting over 65% of total trade (Figure 3). Iron ore (2,601) trade increased until 2010, peaking at 58% with a surge in iron prices, then declined to around 50%. Copper ore (2,603) trade remained steady at approximately 20%. Zinc ore (2,608) trade declined to a stable 5% post-2015, while molybdenum ore (2,613) experienced significant fluctuations, peaking at 6% in 1995 and 8% in 2000, before settling near 2%. Trade in precious metals (2,616) saw overall growth with substantial increases in 2000 and 2020. The trade share of other metal products remained largely consistent with minimal changes.

3.2 Region analysis results

In this study, the global metal minerals trade was categorized into 13 key regions, including North America (NA), South America (SA), Europe (EU), The Commonwealth of Independent States (CIS), the Middle East (MEA), Africa (AF), Oceania (OC), Southeast Asia (SAS), China (CHN), Japan (JPN), Korea (KOR),

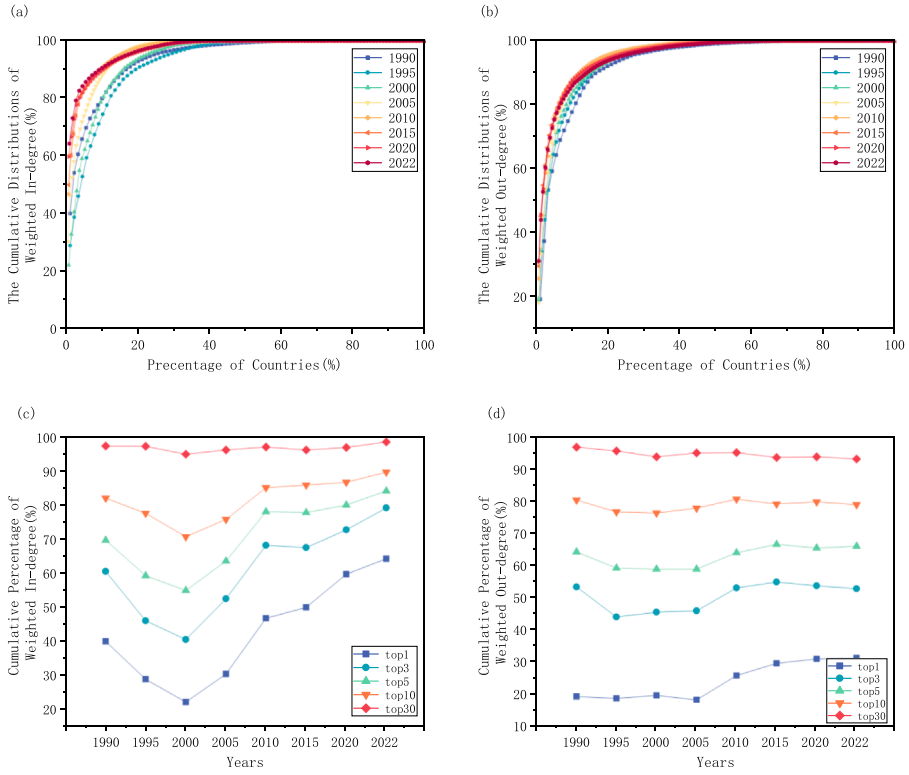


FIGURE 2
The cumulative distributions of the weighted in-degree and weighted out-degree, **(A)** cumulative percentage of weight in-degree, by years; **(B)** cumulative percentage of weight out-degree, by years; **(C)** cumulative percentage of weight in-degree, by top countries; **(D)** cumulative percentage of weight out-degree, by top countries.

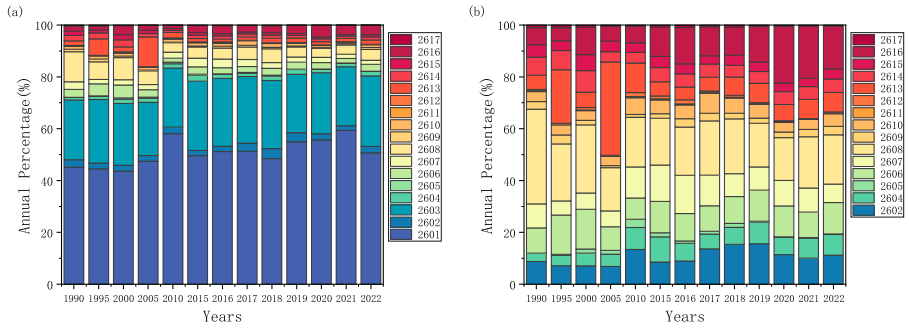


FIGURE 3
Commodity annual percentage, **(A)** included 2,601 and 2,603, **(B)** without 2,601 and 2,603.

and India (IND). Among them, North America, South America, Africa, Oceania, and Southeast Asia were the primary exporters, comprising 88% of total exports in 2022. Meanwhile, Europe, China, Japan, South Korea, and India were the major importers, making up 92% of total imports in 2022. Figure 4 shows the evolution of trade flows, with the left axis indicating trade origins, the right axis showing destinations, line widths signifying trade volumes, and colors denoting distinct regions.

North America has historically been a key player in global metal minerals trade, with robust export capabilities, though its influence has waned in recent years. Between 1990 and 2000, Europe and Japan accounted for ~60% of its exports, but post-2000, exports shifted toward China and South Korea. By 2015, China became the second-largest destination (~28%), while South Korea's share rose to ~10%, matching Japan's. Internal trade remains significant, comprising ~20% of exports and ~40% of imports, with the latter increasing since 2000, reflecting growing regional self-sufficiency.

South America, the largest global exporter, maintained ~30% of global trade volume from 1990 to 2015. However, outdated policies and inefficiencies in major exporters like Chile and Brazil slowed

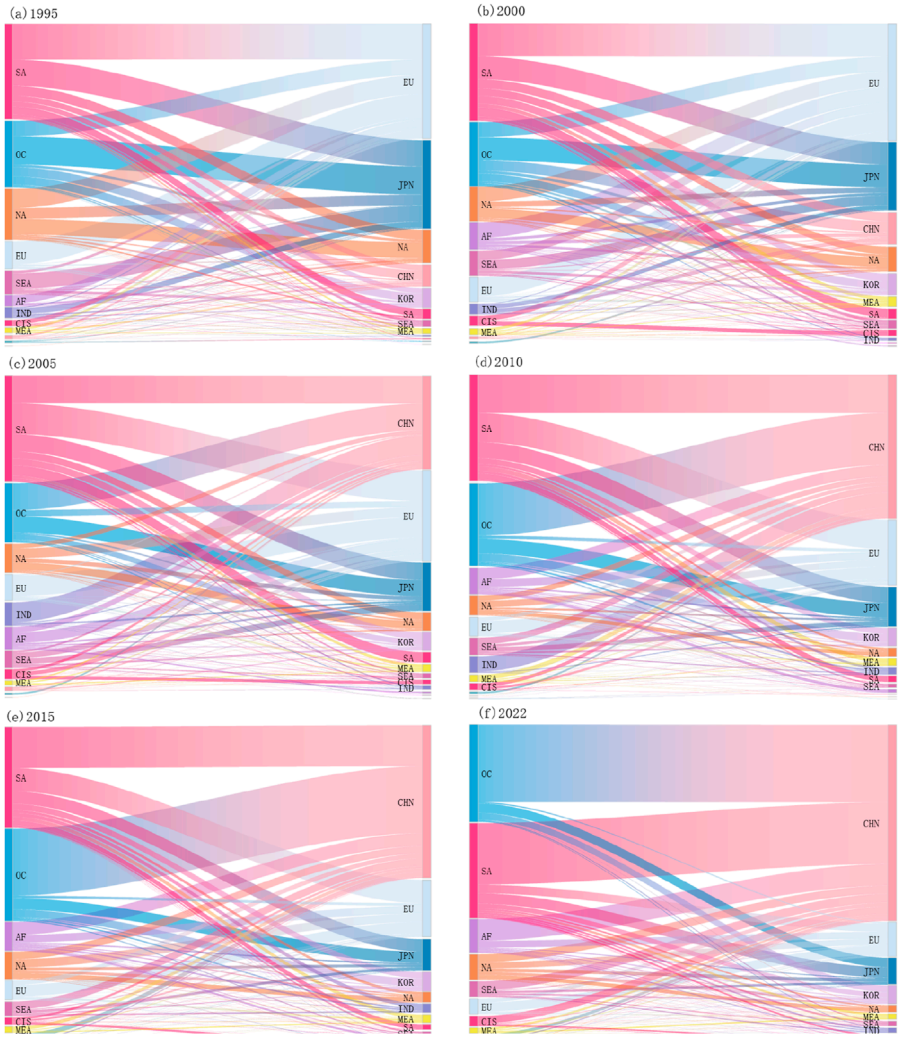


FIGURE 4
The evolution of metal minerals trade flows, by regions. **(A–F)** are the gobal metal minerals trade flows in 1995, 2000, 2005, 2010, 2015, 2022. The regions/countries in the left represent the exporter, and the regions/countries in the right represent the importer.

production growth, leading to its overtaking by Oceania in 2019. Initially, Europe and Japan dominated as export destinations, but China's industrialization shifted trade patterns, with its share rising from 9% (2000) to 65% (2022), marking a transition from multi-regional to concentrated trade.

Oceania, the second-largest exporter, saw its global export share rise from 22% (1995) to 30% (2015), surpassing South America in 2019. Initially, Japan and Europe were primary destinations (~60% combined), but China's demand surged post-2000, reaching 80% of Oceania's exports by 2022. The Oceania-China trade flow became the largest globally, accounting for 24% of total trade in 2022, driven by the 2010 free trade agreement with Australia.

Africa emerged as a key supplier, with exports rising rapidly post-2000. Europe was the primary destination initially, but China's share grew significantly, reaching 65% by 2022, while Europe's stabilized at ~20%. Exploration in Northwest and South Africa fueled this growth, solidifying China's dominance.

Southeast Asia has been a stable exporter, contributing ~7% of global exports (1990–2005), declining slightly to ~5% (2005–2022). Japan and Europe were initially dominant (~70% combined in 2005), but China's share rose sharply, stabilizing at ~55% post-2015, while Japan and Europe's shares fell to 15% and 5%, respectively.

Europe, a major importer, saw its global import share drop from 38% (1995–2000) to 11% (2022). South America, North America, and Africa are key sources, with intra-European trade rising to 26% of imports by 2022. Europe reduced reliance on South America (44%–25%) while increasing imports from North America and Africa.

China became the largest importer, with its global import share surging from 1990. By 2022, it accounted for 64% of global imports ($224.3 billion), driven by industrialization and iron ore demand. Oceania (39%), South America (30%), and Africa (10%) are primary sources, with Oceania's share growing significantly post-2010 due to Australia's iron ore exports.

Japan's import scale declined annually, falling to 9% of global imports by 2022 (from 40% in 1990). South America, Oceania, and North America are key sources, with imports from Southeast Asia dropping to 8% by 2022. Free trade agreements with South America, Oceania, and North America between 2010 and 2015 reshaped its import structure.

South Korea maintains steady imports (~6% of global total), with Oceania and North America becoming dominant sources (~56% combined by 2022) due to its participation in the Trans-Pacific Strategic Economic Partnership Agreement with Australia and the U.S. from 2010 to 2015.

India transitioned from a supplier to a major consumer. Exports peaked at 3.5% of global trade in 2010 but declined due to export tariffs and mining restrictions post-2009, particularly on iron ore. Imports rose during rapid economic growth (2000–2010), stabilizing at ~2% of global imports by 2022. A production resurgence from 2014 helped recover mineral output and exports.

In summary, the global trade of metal minerals is centered around two main trade hubs: Europe, with North America and Africa as suppliers, and the Asia-Pacific, with China, Japan, and South Korea as demand centers and Oceania and Southeast Asia as suppliers. South America serves as a key mineral supplier to both trade circles. From 1990 to 2000, the trade was balanced between the Europe-America and Asia-Pacific circles, with Japan as a central player, leading to a diversified trade pattern. By 2000–2010, Japan's trade diminished, shifting the Asia-Pacific hub to China, sparking competition with the Europe-America trade circle. From 2010 to 2022, China's trade continued to grow, while Europe and Japan's trade stabilized, leading to a decrease in their proportion of the total trade volume. By 2022, China accounted for over 60% of the trade, centralizing the trade circle around it. Overall, the trade pattern evolved from a diversified to a centralized system, increasing the instability of the global trade system. Additionally, India's trade role shifted from supplier to demander due to increased domestic demand for iron ore during its industrialization phase, reducing exports to meet internal needs.

3.3 Communities detection results

In order to further investigate the evolution of trading communities in the global metal minerals trade, this paper constructs a complex network model and employs the map equation framework for community detection and monitoring. The trade flow data representing 80%, 90%, 95%, and 99% of the accumulated trade weight were extracted annually and categorized accordingly. Utilizing the Infomap code base, the data were processed in Python to test community division by year. The analysis revealed that selecting trade data with a 95% cumulative trade weight yielded the most stable and realistic community division. Thus, the paper focuses on countries within this 95% threshold for community analysis.

This study utilized Gephi to visualize the global metal minerals trade network, and the node size indicating a country's weighted degree and line thickness signifying trade value. Due to the large difference between the international situation in 1990 and today, the participating countries in the global metal minerals trade were incomplete. This study focused on the years 1995, 2000, 2005, 2010, 2015, and 2022 as key observation points, with detailed classifications displayed in Figure 5. Communities are divided as follows:

(1) In the early stage, the communities with Japan as the center and India, Malaysia, and other Southeast Asian countries as the main members are marked with orange;
(2) Those with China as the center and South Korea, Indonesia, Thailand, and other Asian countries as the main members in the later period are marked in yellow;
(3) Those with Australia as the center and Brazil, Chile, the United States, Canada and other countries in the Americas as the main members are marked in green;
(4) Communities centered in Germany with major members from Europe and some North African countries, such as Spain, Italy, the United Kingdom and Egypt, are marked in blue;
(5) The community with Russia as the center and Ukraine, Poland, Uzbekistan, Kazakhstan and other CIS countries as the main members, marked with cyan;
(6) Communities with Southern African members centered on South Africa, identified in pink;
(7) A transitory Brazil-centered trading community, marked in purple;
(8) And a trading community that appears briefly and centers on Chile, marked in red.

Between 1990 and 2000, Japan served as the largest global trade hub for metals and minerals. However, its prominence declined after 2005, as it became one of the top five centers within the Australia-Americas community. China, initially part of the Australia-Americas community, experienced rapid growth following its post-2000 economic reforms, emerging as the world's largest metal minerals trader by 2005 and forming a new trade community centered on itself. From 2005 to 2022, this China-led community expanded significantly, becoming the largest trade group. The Australia-Americas community, characterized by abundant mineral resources, has consistently played a pivotal role in global trade, maintaining a relatively stable scale and membership. Countries such as Brazil and Chile temporarily formed a separate community in 2000 but rejoined the Australia-Americas community after 2005. The European community, primarily comprising EU member states and select North African countries, has remained relatively stable. However, between 2010 and 2015, some members temporarily transitioned to other communities before rejoining the German-led European community. The Russian community exhibited gradual growth until 2015 but experienced a contraction in 2022, as some members, including Ukraine, moved to the German European community. The South African community, while smaller in scale, has demonstrated remarkable stability, with its membership remaining largely unchanged over an extended period.

Globally, from 1990 to 2000, multiple trade centers coexisted, creating a competitive trade network pattern. From 2000 to 2010, China's economic boom and Japan's stable demand post-industrialization led to a shift in the trade network, with China and Australia emerging as the main trade centers. From 2010 to 2022, China became a super-demand center in the global metal minerals trade, with Australia as the main supply center, leading to a concentration of trade and a shift from a multi-centered to a dual-centered pattern dominated by China and Australia.

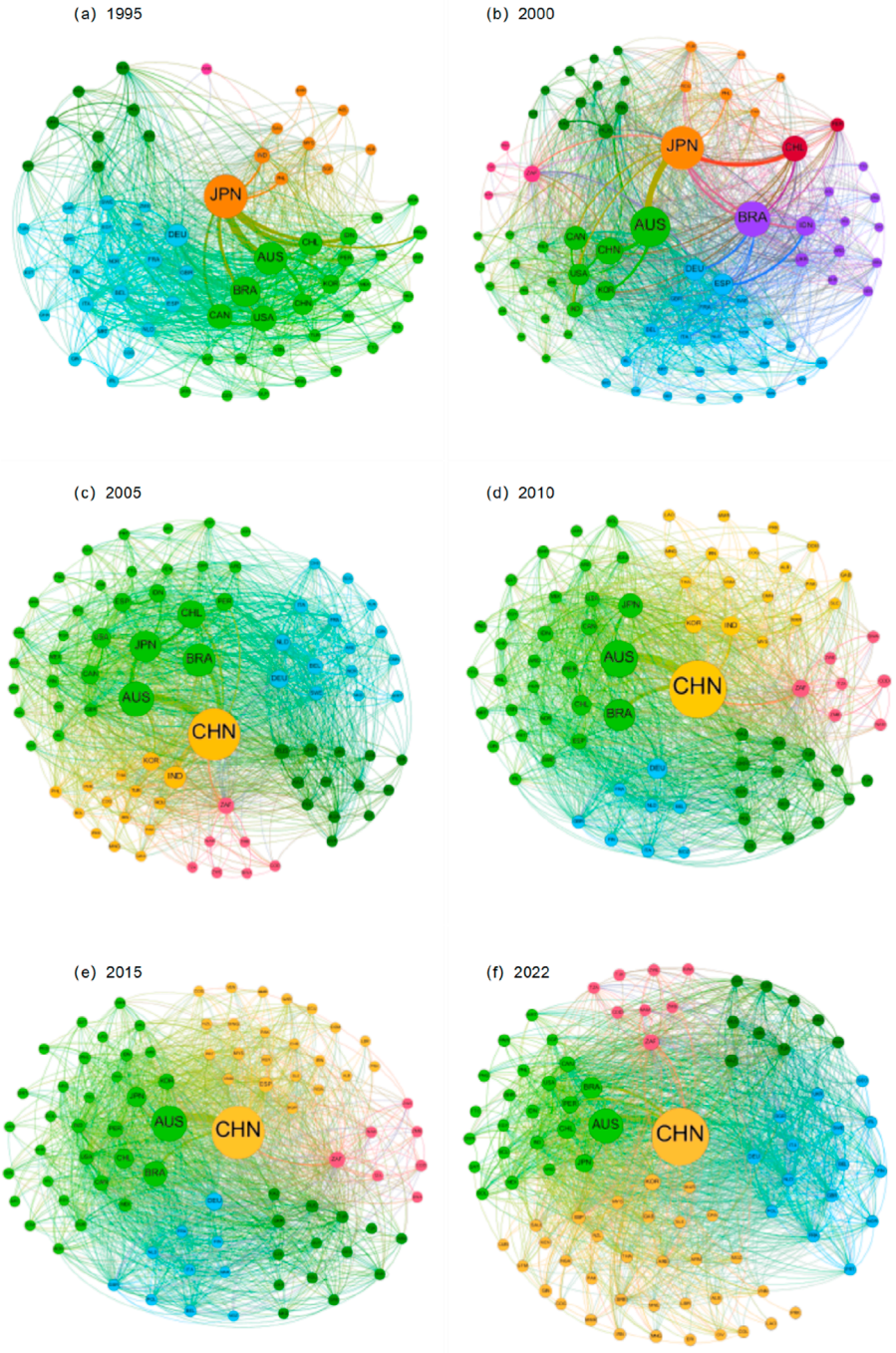


FIGURE 5
The evolution of gobal metal minerals trade community. **(A–F)** show the community division in 1900, 2000, 2005, 2010, 2015,2022. The different colors represent different communities, and the specific division is shown in the Section 3.3.

To further study the key factors of the evolution of the global metal minerals trade structure, this study analyzed the overall situation of the global metal minerals trade, the sub-regional evolution of the global metal minerals trade, and the evolution of the complex network community of the global metal minerals trade, and further collected the characteristic index data of each country in the complex network. And analyzed the leading countries in the numerical rankings.

The weighted in-degree percentage is a key measure of a country's share in global metal and mineral trade imports (Figure 6A), indicating the trade network's import center. Japan's indegree percentage declined significantly from 1990 to 2000, then slowed until 2010, leveling off around 10% of global imports from 2010 to 2022. In contrast, China's in-degree percentage showed gradual growth from 1990 to 2000, a sharp increase between 2000 and 2010, and a slower rise post-2010,

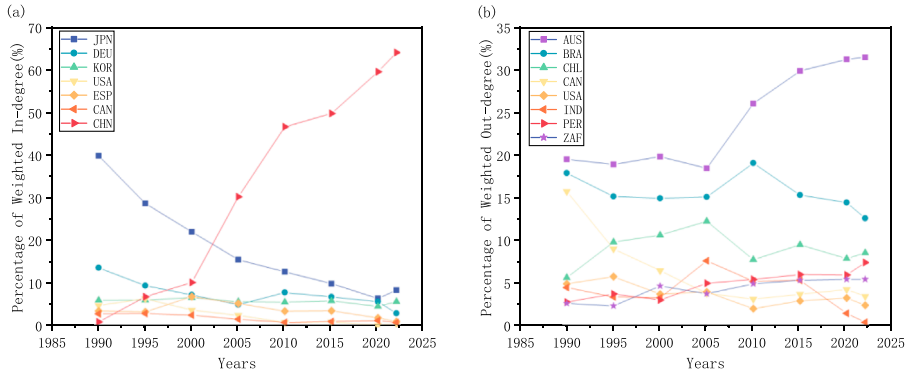


FIGURE 6
Percentage of weighted in-degree and weighted out-degree of major countries, 1990 to 2022. **(A)** Percentage of weighted in-degree of major countries; **(B)** Percentage of weighted out-degree of major countries. Where JPN-Japan, DEU- Germany, KOR- Korea, USA- United States, ESP- Spain, CAN- Canada, CHN-China, AUS- Australia, BRA- Brazil, CHL- Chile, IND- India, PER- Peru, ZAF- Zambia.

eventually stabilizing at a higher share of global imports. Germany's in-degree percentage also decreased, with a marked drop by 2000, followed by stabilization at around 8%. Meanwhile, countries such as the United States, South Korea, Spain, and Canada maintained relatively stable in-degree percentages throughout 1990–2022. Over this period, the trade network's import hub shifted from Japan to China.

The weighted out-degree percentage is a crucial measure of a country's contribution to global metal and mineral trade exports (Figure 6B). Historically, Australia has been the top exporter with a stable outdegree percentage of around 20% from 1990 to 2005, followed by rapid growth to 2015, and stabilization thereafter. Australia has consistently been the trade's export hub from 1990 to 2022. Brazil, the second-largest exporter, maintained a roughly 15% share, with a spike in 2010 due to Chile's decreased exports. Chile's outdegree rose until 2005, then stabilized at 8%. Canada experienced a decline in 2005, followed by stabilization. The U.S. showed a steady decline, while South Africa trended gently upward, both achieving stability. Overall, Australia remains the central exporter, with other countries maintaining a steady proportionate structure in the global trade network.

Betweenness centrality evaluates a nation's ability to influence trade network dynamics (Figure 7A). In 1990, Germany exhibited the highest betweenness centrality, reflecting significant network influence, although this influence fluctuated and diminished between 1995 and 2022. The United States maintained a high centrality index from 1995 to 2005, demonstrating stable control over global trade, but experienced a gradual decline after 2005 while retaining considerable international influence. China showed a sharp increase in betweenness centrality from 1990 to 2010, signaling its growing prominence in the global metal and mineral trade. It has remained the top-ranked country from 2010 to 2022, despite a slight decrease in its centrality value. Canada exhibited an upward trend until 2000, followed by a gradual decline, eventually stabilizing at a moderate level of influence. India's centrality increased until 2010 before declining, while South Africa showed a slow but steady decline, maintaining a mid-tier position. Spain, by contrast, displayed consistent stability throughout the period.

Closeness centrality indicates a node's proximity to others, reflecting a country's capacity to exert counter-influence within a network (Figure 7B). Germany's closeness centrality declined from 1990 to 2010 but rebounded by 2022, indicating a recovery in counter-control ability. The US maintained a high and stable closeness centrality from 1995 to 2022, demonstrating strong counter-control and trade security, with extensive global trade ties and high trade freedom. China's centrality was initially low but rose to the top after 2010, showing improved trade security and a continued upward trend to 2022. South Africa, with rich mineral resources and self-sufficiency, has a consistently high and stable closeness index, reflecting its stable counter-control ability. Australia, similarly resource-rich, has a lower overall closeness index. India's index rose steadily, enhancing its counter-influence in global metal minerals trade, though still lagging behind top countries. Brazil's closeness index followed a pattern of decline and recovery, with a gradual strengthening of counter-control after hitting a low in 2005.

According to the analysis of the value change of each node characteristic index, the study finds that Japan, China, Australia, Brazil, the United States, and some European countries play an important role in the structure evolution of the global metal minerals trade network. Japan, despite a declining import share from 1990 to 2022 and weak control in the trade network due to limited domestic resources, remains a key player. China, as the current trade network's hub and import center, has closely intertwined development with the overall metal minerals trade from 1990 to 2022, enhancing its counter-control to safeguard trade security. Australia, the second-largest trading entity and export center, maintains a stable trade relationship with China and ensures its trade security through stable counter-control. Brazil, with a consistent export share and rich metal resources, is a major exporter and counter-control player. The United States, while not the center of imports or exports, exerts significant control and offers stability as a crucial trade network connector between countries.

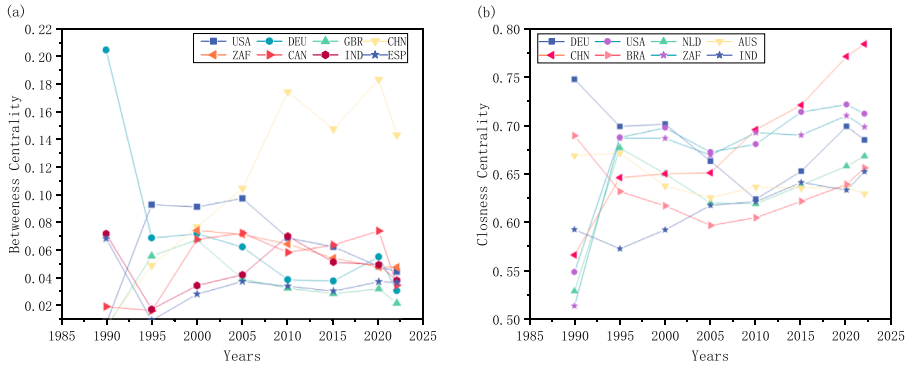


FIGURE 7
Betweenness centrality and closeness centrality of major countries, 1990 to 2022. **(A)** Betweenness centrality of major countries; **(B)** Closeness centrality of major countries. Where USA- United States, DEU- Germany, GBR- United Kingdom, CHN-China, ZAF- Zambia, CAN- Canada, IND- India, ESP- Spain, NLD- Netherlands , AUS- Australia, BRA- Brazil.

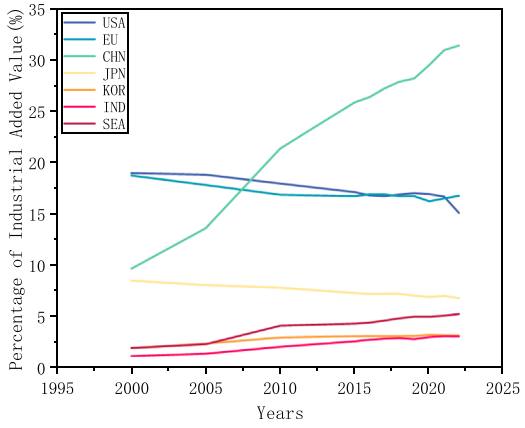


FIGURE 8
Percentage of industrial added value of major countries and regions, 2000–2022.

4 Discussion

4.1 Metal minerals and industrial development

Metal minerals are crucial for societal industrial development, with demand often rising alongside industrialization. The range of metal minerals needed for comprehensive industrial development is vast, and their distribution is typically uneven, meaning few nations can rely solely on domestic production. This creates a significant link between the scale of metal minerals trade and industrial development. Figure 8 shows that since 2000, the manufacturing industry's value-added has declined in the U.S., Europe, and Japan, China's rapid growth has correspondingly increased its import position in the global metal minerals trade.

At present, the gap between the trade weight of other centrosomes and super centrosomes is wide, except for China and Australia, other countries' trade changes are not obvious, but China 2021 proposed the "14th Five-Year Plan" [40], the domestic launch of the new steel policy, steel production has declined, iron ore demand growth will be limited. India has had the world's highest GDP growth rate in recent years, the future steel demand will further increase, while domestic iron ore exports will also reduce, further to the demand-side role transformation. At the same time, investment in African mineral exploration projects is still advancing, and the production capacity of metal minerals in Africa will be further improved in the future [41]. In the future, the global metal minerals trade structure may change again, from the structure of China as the supercenter to the multi-center development of the trade structure, while Africa and India will become the new supply and demand centers, the global metal minerals trade structure from a highly concentrated state will once again turn to a diversified trade pattern.

4.2 Policy implications and resource security

Mineral resource trade is a crucial component of the metal minerals supply chain, and over-reliance on concentrated import sources can jeopardize a nation's resource security. As the trade network becomes more intricate, nations have implemented policies to diversify trade and safeguard domestic resources. For instance, the U.S. formed the Mineral Security Partnership [42] with allies on June 14, 2022, aiming to curtail reliance on foreign adversaries for critical minerals and ensure national and economic security [43]. Concurrently, the European Union introduced the Critical Raw Materials Act on September 14, 2022, mandating that by 2030, reliance on a single third country for any strategic material should not surpass 65% of the EU's annual consumption at any stage of processing [44]. These policies indicate a future trend where global metal minerals trade, under the influence of trade diversification, will evolve towards alliance-based structures, with developed nations and their allies forming closer trade communities, fostering a competitive environment with supercenters.

Indonesia, Chile, Peru, the Democratic Republic of Congo, Australia, and Canada, as major suppliers of metal minerals, have implemented policies to curb the direct export of raw minerals

and bolster their economies. Indonesia's 2014 ban on raw mineral exports, including nickel ore and bauxite, mandates local smelting or refining, attracting foreign investment in downstream industries and increasing export value [45]. Chile, Peru, and the DRC have raised export taxes on minerals like copper, lithium, and cobalt, and enforced national participation in mining to boost economic revenue [46, 47]. Australia and Canada promote domestic mining through key mineral lists, aiming to enhance economic growth, competitiveness, and employment [48, 49]. While China remains a significant export destination for Australian minerals, potential policy shifts influenced by the U.S. and its allies could encourage Australia to explore new markets in emerging economies like India and Southeast Asia, driving global trade structure evolution in metal minerals.

At the same time, the global energy transition is driving changes in the consumption structure of metal minerals. In the future, the energy consumption pattern dominated by clean energy such as wind power, photovoltaic, and hydrogen energy will greatly increase the demand for new energy metals such as copper, aluminum, nickel, cobalt, and lithium while the proportion of demand for iron will be reduced [50]. Moreover, the new energy sector also requires a more diverse range of metal minerals, particularly scarce minor metals such as gallium, germanium, indium, tellurium, and antimony, which will see a significant rise in demand, leading to a more diversified trade in metal minerals varieties globally [51].

4.3 Impact of geopolitical events on metal minerals trade

Geopolitical events have had profound impacts on the global trade structure of metal minerals. Trade frictions and conflicts have led to short-term disruptions and long-term shifts in trade patterns. In the short term, geopolitical events can cause significant disruptions in supply chains, leading to supply shortages and price increases. In the long term, these events drive structural changes in trade networks. For example, The ongoing Russia-Ukraine conflict has hit the European key mineral supply chain. As a major global producer of key minerals such as nickel and palladium, the instability of Russia's supply has prompted European companies to accelerate the search for alternative resources [52]. Such as Finland has reached cooperation with Canada and increased investment in the development of nickel ore resources in Canada to make up for the potential gap caused by changes in Russia's supply. In the Middle East, the tense geopolitical situation have put the transportation of minerals passing through the region at potential risk. The conflicts have triggered concerns about supply disruptions. In response to this situation, many major Asian electronics - producing countries, such as Japan and South Korea, have strengthened cooperation with Australia and Brazil to ensure the stable import of key minerals such as rare earths, gallium, and germanium [53].

5 Conclusion

Combining resource endowment, national policy, industrialization process, energy transition, and other influencing factors, this study analyzes the evolution process and evolution characteristics of the global metal minerals trade structure from 1990 to 2022 by region, community, and country. The main conclusions of this paper are as follows:

(1) From 1990 to 2022, the global trade in metals and minerals shows an overall growth trend, which can be divided into three stages: A. Slow development period (1990–2000), B. Rapid growth period (2000–2010), C. Slow growth period (2010~2022). With the development of international trade, an increasing number of countries have engaged in trade activities, while the concentration of metal minerals imports has gradually shifted toward the top 10 importing countries. A small number of nations dominate the majority of global metal minerals trade. Currently, iron and copper account for over 65% of the total trade volume, making them the most significant traded metal products. However, with the rise of the new energy sector, the demand for new energy metals is expected to rise, while the demand for traditional industry metals will diminish, leading to a more diversified trade in metal minerals.

(2) Great changes have taken place in the structure of global metal minerals trade from 1990 to 2022. In the first stage, competition primarily occurred between the Japan-centered trade circle and the Europe-centered trade circle. As Japan entered the later stages of industrialization, its demand for metal minerals stabilized. In the second stage, the structure gradually shifted to competition between the Europe-centered trade circle and the China-centered trade circle. In the third stage, after 2015, China's demand for metal minerals continued to grow, and the global trade structure of metal minerals gradually evolved into a dual trade structure with China as the super demand center and Australia as the super supply center. On the whole, it shows an evolution from decentralization to centralization.

(3) Through the analysis of the numerical changes of each node characteristic index, the research finds that Japan, China, Australia, Brazil, the United States, and some European countries play an important role in the evolution of the structure of the global metal minerals trade network. Among countries with high import weight in metal minerals trade, those with higher betweenness centrality tend to have greater influence compared to nations with lower import weight. This indicates that the structure of the global metal minerals trade network is primarily shaped by demand centers. Conversely, countries with abundant domestic metal minerals resources exhibit relatively higher closeness centrality, suggesting that international metal trade has a comparatively smaller impact on their resource supply.

(4) The global metal minerals trade structure is poised for a shift towards a diversified network with multiple trade centers. This evolution is partly due to China's peak demand for iron ore and its commitment to carbon neutrality, which will likely reduce foreign iron ore demand and increase the use of recycled steel. Meanwhile, the demand for copper, crucial for the new energy sector, will rise, enhancing trade with copper-rich nations like Chile and Peru. Additionally, emerging economies such as India and Southeast Asia are expected to grow their industrialization, potentially becoming

new demand hubs. Developed countries like the U.S. and Europe are also forging diversified supply chains, fostering a multipolar trade landscape.

It should be noted that this study has examined metal minerals trade data only within HS codes 2,601 to 2,617 from UN Comtrade, yet it acknowledges limitations as spodumene and lithium carbonate trade data with 8-bit codes which do not open publish date, and the intermediate products of cobalt metallurgy by wet process (HS code: 8105201000) which dominated the cobalt trade is not included. The research could be enhanced by incorporating these elements. Additionally, the study applies a disjoint community rule in network analysis, which may overlook the multi-community influence of key nodes, such as bridge countries, that are instead absorbed into larger communities by more dominant nodes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YZ: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Methodology, Project administration, Software, Visualization, Writing–original draft, Writing–review and editing. QC: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Project administration, Writing–original draft, Writing–review and editing. KW: Methodology, Writing–review and editing. QG: Methodology, Writing–review and editing. JX: Project administration, Supervision, Writing–review and editing. TL: Project administration, Supervision, Writing–review and editing. GZ: Project administration, Supervision, Writing–original draft. ML: Project administration, Supervision, Writing–original draft. CS: Formal Analysis, Writing–review and editing. JM: Supervision, Methodology, Writing–review and editing, Writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by Deep Earth Probe and Mineral Resources Exploration - National Science and Technology Major Project [2024ZD1002001]; National Natural Science Foundation of China (NSFC) [grant numbers 42271281]; China Geological Survey, Ministry of Natural Resources [grant number DD20230040; DD20221694; DD20211405].

Acknowledgments

We are very grateful to the academic suggestions provided by Zhang Hongwei of Central South University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Song Y, Bai W, Zhang Y. Resilience assessment of trade network in copper industry chain and the risk resistance capacity of core countries: based on complex network. *Resour Pol* (2024) 92:105034. doi:10.1016/j.resourpol.2024.105034

2. Chen W, Zhang J, Yu Z, Zhao X. Structure and evolution of global lead trade network: an industrial chain perspective. *Resour Pol* (2024) 90:104735. doi:10.1016/j.resourpol.2024.104735

3. Kang X, Wang M, Wang T, Luo F, Lin J, Li X. Trade trends and competition intensity of international copper flow based on complex network: from the perspective of industry chain. *Resour Pol* (2022) 79:103060. doi:10.1016/j.resourpol.2022.103060

4. Li B, Li H, Dong Z, Lu Y, Liu N, Hao X. The global copper material trade network and risk evaluation: a industry chain perspective. *Resour Pol* (2021) 74:102275. doi:10.1016/j.resourpol.2021.102275

5. Hu X, Wang C, Zhu X, Yao C, Ghadimi P. Trade structure and risk transmission in the international automotive Li-ion batteries trade. *Resour Conservation Recycling* (2021) 170:105591. doi:10.1016/j.resconrec.2021.105591

6. Ma Y, Wang M, Li X. Analysis of the characteristics and stability of the global complex nickel ore trade network. *Resour Pol* (2022) 79:103089. doi:10.1016/j.resourpol.2022.103089

7. Shuxian Z, Xuanru Z, Pei Z, Wanli X, Yawen H, Hongchang H, et al. Impact of countries' role on trade prices from a nickel chain perspective: based on complex network and panel regression analysis. *Resour Pol* (2022) 78:102930. doi:10.1016/j.resourpol.2022.102930

8. Shi Q, Sun X, Xu M, Wang M. The multiplex network structure of global cobalt industry chain. *Resour Pol* (2022) 76:102555. doi:10.1016/j.resourpol.2022.102555

9. Li Y, Huang J, Zhang H. The impact of country risks on cobalt trade patterns from the perspective of the industrial chain. *Resour Pol* (2022) 77:102641. doi:10.1016/j.resourpol.2022.102641

10. Sun X, Shi Q, Hao X. Supply crisis propagation in the global cobalt trade network. *Resour Conservation Recycling* (2022) 179:106035. doi:10.1016/j.resconrec.2021.106035

11. Liu M, Li H, Zhou J, Feng S, Wang Y, Wang X. Analysis of material flow among multiple phases of cobalt industrial chain based on a complex network. *Resour Pol* (2022) 77:102691. doi:10.1016/j.resourpol.2022.102691

12. Zhu M, Zhou X, Zhang H, Wang L, Sun H. International trade evolution and competition prediction of boron ore: based on complex network and link prediction. *Resour Pol* (2023) 82:103542. doi:10.1016/j.resourpol.2023.103542

13. Hao X, An H. Comparative study on transmission mechanism of supply shortage risk in the international trade of iron ore, pig iron and crude steel. *Resour Pol* (2022) 79:103022. doi:10.1016/j.resourpol.2022.103022

14. Sauer PC, Seuring S. Extending the reach of multi-tier sustainable supply chain management – insights from mineral supply chains. *Int J Prod Econ* (2019) 217:31–43. doi:10.1016/j.ijpe.2018.05.030

15. Ren B, Li H, Wang X, Shi J, Ma N, Qi Y. The flow of embodied minerals between China's provinces and the world: a nested supply chain network perspective. *Resour Pol* (2022) 78:102853. doi:10.1016/j.resourpol.2022.102853

16. Chen Z, Wang H, Liu X, Wang Z, Wen S. Risk diffusion of international oil trade cuts: a network-based dynamics model. *Energy Rep* (2022) 8:11320–33. doi:10.1016/j.egyr.2022.08.244

17. Hao H, Xing W, Wang A, Song H, Han Y, Zhao P, et al. Multi-layer networks research on analyzing supply risk transmission of lithium industry chain. *Resour Pol* (2022) 79:102933. doi:10.1016/j.resourpol.2022.102933

18. Khawaja FR, Zhang Z, Memon Y, Ullah A. Exploring community detection methods and their diverse applications in complex networks: a comprehensive review. *Social Netw Anal Mining* (2024) 14(1):115. doi:10.1007/s13278-024-01274-1

19. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett* (2001) 86(14):3200–3. doi:10.1103/PhysRevLett.86.3200

20. Yang W, Chakrabarti D, Chenxi W, Faloutsos C. Epidemic spreading in real networks: an eigenvalue viewpoint. In: 22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings; 06-08 October 2003; Florence, Italy. IEEE (2003). p. 25–34. doi:10.1109/RELDIS.2003.1238052

21. Wu L, Waber B, Aral S, Brynjolfsson E, Pentland S. *Mining face-to-face interaction networks using sociometric badges: evidence predicting productivity in IT configuration* (2008). Available from: http://dx.doi.org/10.2139/ssrn.1130251.

22. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* (2005) 438(7066):355–9. doi:10.1038/nature04153

23. Barabási A-L. The origin of bursts and heavy tails in human dynamics. *Nature* (2005) 435(7039):207–11. doi:10.1038/nature03459

24. Zuo Z, McLellan BC, Li Y, Guo H, Cheng J. Evolution and insights into the network and pattern of the rare earths trade from an industry chain perspective. *Resour Pol* (2022) 78:102912. doi:10.1016/j.resourpol.2022.102912

25. Xu-Guang Z, Xue-Hong Z, Jin-Yu C. Trade dependence network structure of tantalum trade goods and its effect on trade prices: an industry chain perspective. *Resour Pol* (2022) 79:103065. doi:10.1016/j.resourpol.2022.103065

26. Zhang H, Wang X, Tang J, Guo Y. The impact of international rare earth trade competition on global value chain upgrading from the industrial chain perspective. *Ecol Econ* (2022) 198:107472. doi:10.1016/j.ecolecon.2022.107472

27. Huang J, Ding Q, Wang Y, Hong H, Zhang H. The evolution and influencing factors of international tungsten competition from the industrial chain perspective. *Resour Pol* (2021) 73:102185. doi:10.1016/j.resourpol.2021.102185

28. Blöcker C, Nieves JC, Rosvall M. Map equation centrality: community-aware centrality based on the map equation. *Appl Netw Sci* (2022) 7(1):56. doi:10.1007/s41109-022-00477-9

29. Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci* (2008) 105(4):1118–23. doi:10.1073/pnas.0706851105

30. Edler D, Bohlin L, Rosvall M. Mapping higher-order network flows in memory and multilayer networks with Infomap. *Algorithms* (2017) 10(4):112. doi:10.3390/a10040112

31. Rosvall M, Axelsson D, Bergstrom CT. The map equation. *The Eur Phys J Spec Top* (2009) 178(1):13–23. doi:10.1140/epjst/e2010-01179-1

32. Bohlin L, Edler D, Lancichinetti A, Rosvall M. Community detection and visualization of networks with the map equation framework. In: *Measuring scholarly impact*. Springer (2014). p. 3–34. doi:10.1007/978-3-319-10377-8_1

33. United Nations. *UN comtrade* (2024). Available from: https://comtradeplus.un.org/(Accessed July 22, 2024).

34. Newman MEJ. *Networks: an introduction*. Oxford University Press (2010).

35. Barabási A-L. *Network science*. Cambridge University Press (2016).

36. Serrano MÁ, Boguñá M. Topology of the world trade web. *Phys Rev E* (2003) 68(1):015101. doi:10.1103/PhysRevE.68.015101

37. Golbeck J. Network structure and measures. In: *Analyzing the social web*. Elsevier (2013). p. 25–44. doi:10.1016/B978-0-12-405531-5.00003-1

38. Bhattacharya S, Sinha S, Dey P, Saha A, Chowdhury C, Roy S. Online social-network sensing models. In: *Computational intelligence applications for text and sentiment data analysis*. Elsevier (2023). p. 113–40. doi:10.1016/B978-0-32-390535-0.00010-0

39. Hansen DL, Shneiderman B, Smith MA, Himelboim I. Calculating and visualizing network metrics. In: *Analyzing social media networks with NodeXL*. Elsevier (2020). p. 79–94. doi:10.1016/B978-0-12-817756-3.00006-6

40. Xinhua News Agency. *People's Republic of China outline of the 14th five-year plan for national economic and social development and the long-range objectives through the year 2035*. Xinhua News Agency (2021). p. 79–126.

41. Wang QX, Shi JP, Yang RH, Yuan CH, Lian CY. The situation of solid mineral exploration in Africa and its influence on China's "going out" strategy. *China Mining* (2022) 31:7–14. doi:10.12075/j.issn.1004-4051.2022.04.009

42. US Department of State. *Minerals security partnership*. US Department of State (2023). Available from: https://www.state.gov/minerals-security-partnership/(Accessed July 22, 2024).

43. Building Resilient Supply Chains. *Revitalizing American manufacturing, and fostering broad-based growth the white house* (2021). p. 1–250. Available from: https://whitehouse.gov/wp-content/uploads/2021/06/100-day-supply-chain-review-report.pdf?ref=hackernoon.com (Accessed July 22, 2024).

44. European Commission. *Critical raw materials Act*. European Commission (2022). Available from: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/green-deal-industrial-plan/european-critical-raw-materials-act_en#:~:text=The%20European%20Critical%20Raw%20Materials%20Act%20is%20an,supply%20of%20critical%20raw%20materials%20in%20the%20EU (Accessed July 22, 2024).

45. Indonesian Government *Prohibition of the export of nickel ore*. Indonesian Government (2020). Available from: https://www.iea.org/policies/16084-prohibition-of-the-export-of-nickel-ore (Accessed July 22 2024)

46. Lagos G. Mining nationalization and privatization in Peru and in Chile. *Mineral Econ* (2018) 31(1–2):127–39. doi:10.1007/s13563-017-0124-9

47. Bloomberg. *Chile a step closer to nationalizing copper and lithium*. Bloomberg (2022). Available from: https://www.mining.com/chile-a-step-closer-to-nationalizing-copper-and-lithium/(Accessed July 22, 2024).

48. Geoscience Australia. *Critical minerals strategy 2023–2030*. Geoscience Australia (2023). Available from: https://www.industry.gov.au/publications/critical-minerals-strategy-2023-2030 (Accessed July 22, 2024).

49. Government of Canada. *Canada's critical minerals*. Government of Canada (2024). Available from: https://www.canada.ca/en/campaign/critical-minerals-in-canada/critical-minerals-an-opportunity-for-canada.html (Accessed July 22, 2024).

50. Li C, Li X, Zhang M, Yang B. Sustainable development through clean energy: the role of mineral resources in promoting access to clean electricity. *Resour Pol* (2024) 90:104675. doi:10.1016/j.resourpol.2024.104675

51. D'Orazio P. Assessing the fiscal implications of changes in critical minerals' demand in the low-carbon energy transition. *Appl Energy* (2024) 376:124197. doi:10.1016/j.apenergy.2024.124197

52. Khurshid A, Chen Y, Rauf A, Khan K. Critical metals in uncertainty: how Russia-Ukraine conflict drives their prices? *Resour Pol* (2023) 85:104000. doi:10.1016/j.resourpol.2023.104000

53. Zhao P, Tong Z. The relationships between geopolitics and global critical minerals shipping: a literature review. *Ocean and Coastal Management* (2025) 262:107559. doi:10.1016/j.ocecoaman.2025.107559

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 27 March 2025
DOI 10.3389/fphy.2025.1559283

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Chan Liu,
China University of Mining and Technology, Beijing, China
Xiaowen Shi,
Beijing Normal University, China

*CORRESPONDENCE
Wei Cui,
cuiw@cugb.edu.cn

RECEIVED 12 January 2025
ACCEPTED 13 March 2025
PUBLISHED 27 March 2025

CITATION
Zhang Y and Cui W (2025) Research on characterization and prediction of bond risk factors based on machine learning: evidence from the China.
Front. Phys. 13:1559283.
doi: 10.3389/fphy.2025.1559283

COPYRIGHT
© 2025 Zhang and Cui. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Research on characterization and prediction of bond risk factors based on machine learning: evidence from the China

Yaowen Zhang and Wei Cui*

School of Economics and Management, China University of Geosciences, Beijing, China

Introduction: The scale of default on credit bonds in China has been expanding. Credit bond defaults not only increase the financing costs of enterprises but also affect the efficiency of debt issuance and even lead to the spread of risks in the financial market. Accurately identifying bond default risks, clarifying the characteristics of bond defaults, and understanding the default risk mechanism are of crucial importance.

Methods: This paper takes corporate credit bonds as the research object and analyzes bond defaults from both macro and micro perspectives. From a macro perspective, it confirms the logical transmission between macro factors and bond defaults through causal relationships and grasps the overall characteristics of bond defaults by combining association rule mining and descriptive statistical research methods. Bonds are divided into a risk-free bond group and a risky bond group, and association rules are mined in four dimensions: the bond issuance region of the enterprise, whether the issuer is listed, the attributes of the issuing enterprise, and whether the enterprise bond is guaranteed. Based on these rules, a cross-analysis of bond risk factors is conducted. From a micro perspective, taking each bond as the research object, a bond default identification system is established, and default predictions are made based on the ensemble learning algorithm. The important characteristics of default bonds are analyzed from the perspective of whether the issuer is a state-owned enterprise, and further cause difference analysis is conducted.

Results: The results show that M1 and M2 have an impact on bond defaults, and the ensemble machine learning algorithm can accurately predict bond default risks and obtain key factors for bond risk identification. It is reasonable to choose macro indicators to predict bond defaults.

Discussion: Based on the experimental conclusions, this paper discusses and analyzes the bond risk evolution process and the reasons for risk concentration in certain industries, which is helpful for a comprehensive understanding of bond default risks. Our research can provide tool references and guidance for risk management in the actual bond market.

KEYWORDS

Chinese bond market, bond credit default prediction, macro factors and default causality, combinatorial machine learning, granger causality test, association rule mining

1 Introduction

The bond market is an important component of the capital market. The effective operation of the bond market not only provides financing convenience for market participants but also promptly reflects the operational status of the real economy.

Chinese corporate bond issuance, which began in 1984, started late but has grown rapidly to become the world's second largest bond market. Investors know little about the bond market compared to other forms of financing. The bond market is very sound and low risk for most investors. For a long time, Chinese credit bonds have been regarded as risk-free investment tools by investors, but with the improvement of the marketization degree of Chinese bonds, bond defaults began to appear, and the first case of corporate bond defaults occurred in 2014. Chinese corporate bonds have had a 10-year default history since November 2014, when the first case of "Chaori bonds" actually defaulted. At present, defaults are gradually becoming normal, credit risks are spreading across industries, and market players have stopped unreasonable "risk-free" expectations for Chinese corporate bonds.

By the end of 2023, the total value of outstanding bonds in China's bond market had grown to 155.75 trillion yuan. Starting from 2018, the market witnessed a significant rise in bond defaults. The year 2019 marked the highest number of defaults, with 207 bonds failing to meet their obligations, totaling 166.187 billion yuan. Although the number of defaults dropped slightly in 2020, the total value of defaulted bonds hit its peak at 188.462 billion yuan that year. Following this period, while the overall scale of defaults has shown some reduction, even high-credit-rated companies and centrally-administered state-owned enterprises—typically considered lower-risk entities—have also faced default incidents. This development is likely to have a more profound impact on the credit bond market environment. Bond default will lead to capital loss, market risk intensification, credit risk spread, deterioration of credit environment, reduce investors' investment participation in the bond market, and even lead to systemic financial risk. Therefore, it is very important to analyze and forecast the credit risk factors of bonds. China's corporate bond defaults deserve attention.

As for the prediction methods of bond default, the earliest statistical and quantitative methods are Z-score model [1] and KMV model [2], etc. The statistical and quantitative methods of bond default are developed on the improvement of these models. In recent years, with the development of big data, machine learning algorithms have been applied in the field of credit prediction. Bao et al. [3] found that compared with the majority of existing studies that mainly aim to explain sample fraud and emphasize causal reasoning, the predictive model constructed by ensemble learning can predict accounting fraud more accurately. In the establishment of the bond default system, the current research generally selects the macro environment of the enterprise directly as the predictive index of the machine learning algorithm, while ignoring the causal relationship between the selection of macro indicators and bond default. Therefore, this study integrates causal inference with machine learning techniques to systematically analyze the causal relationships between macroeconomic indicators and bond defaults. It employs machine learning algorithms for predicting bond default risks and validates the effectiveness of the selected macroeconomic indicators through feature importance analysis.

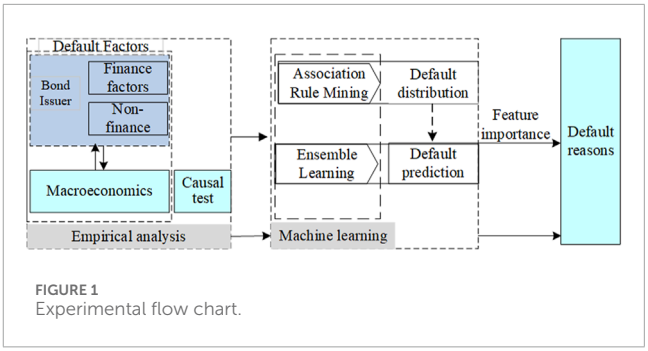


FIGURE 1
Experimental flow chart.

TABLE 1 Confusion matrix.

Actual	Prediction	
	0 (non-risk)	1 (risk)
0 (non-risk)	TN	FP
1 (risk)	FN	TP

The bond market is a crucial link in the transmission of monetary policy. China is in the process of interest rate liberalization. Grasping the risks in China's bond market and clarifying the causal relationship between monetary and financial variables, macroeconomic variables and bond defaults is of great significance.

This study focuses on credit bonds issued by Chinese companies from 2014 to 2022 as the research sample, encompassing corporate bonds, enterprise bonds, medium-term notes, and short-term financing bonds. Bonds that are in default or have been extended are regarded as risky bonds, while bonds that mature normally are considered risk-free bonds. By using causal inference to analyze the causal relationship between macroeconomic indicators and bond defaults, and at the same time providing a reference for the selection of the timing of predictive indicators, machine learning techniques are utilized to predict bond default risks, and the effectiveness of the selected macroeconomic indicators is verified through feature importance analysis. The main experiments are as follows [1]: Conduct an empirical analysis on the impact of macroeconomic factors such as macro monetary policies and the economy on bond defaults through causal inference to explore whether there is a causal relationship between macroeconomic factors and bond defaults [2]. On the basis of descriptive statistics and association rule mining of the attributes of bonds themselves, find the associated features of bond defaults [3]. Process and construct the risk indicator system, and build a bond default risk early warning model through machine learning algorithms. The main innovations and contributions of this paper are as follows [1]: In the selection of indicators, it empirically proves the causal relationship between macroeconomic indicators and bond default [2]; Association rule mining is applied to analyze the relevant characteristics of risk bonds, and the distribution characteristics and rules of risk bonds are obtained [3]; Based on the

TABLE 2 Optimal parameters.

Algorithm	Parameters	Range of values	Optimal parameters
Decision Tree	Criterion	Gini, entropy	Gini
	Max_depth	3,5,10,15,20,25,30	15
	Min_samples_leaf	[1,2,3,4,5]	1
	Min_samples_split	[2,4,6,8,10]	2
RandomForest	N_estimators	[50,1000], step = 1	71
	Max_depth	3,6,9,12,15,18,21,24,27,30	18
GradientBoostingClassifier	Learning_rate	[0.0001-0.1]	0.1
	N_estimators	[50,1000], step = 50	150
	Max_depth	[3,30], step = 1	21
	Max_features	[10,50], step = 1	33
	Subsample	[0.1,0.9], step = 0.1	0.8
	Min_samples_leaf	[60,100], step = 1	70
	Min_samples_split	[50,200], step = 1	100
Catboost	Learning_rate	[0.0001-0.05], step = 0.001	0.136
	Rsm	[0.6, 1.0], step = 0.1	0.60
	Subsample	[0.6, 1.0], step = 0.001	0.736
	Iterations	[20, 50],step = 1	50
	Depth	[3,6], step = 1	4

results of association rules, the combination learning model is used to make group prediction and factor analysis of bond credit status, so as to grasp the differences, and prove that the prediction accuracy of combination learning on bond default is improved compared with the single classification model.

2 Literature reviews

Bond risk identification is the prediction and assessment of the credit status of bond issuance. The following reviews the existing research results from three aspects: the selection of indicators for bond default prediction, including financial and non-financial indicators; the risk transmission between monetary policy and financial markets and causal analysis and machine learning.

In terms of the selection of financial indicators, in 1968, the American economist Beaver [4] was the first to use financial analysis indicators to predict the default of corporate bonds and credit, and to create a univariate financial early warning model with the idea of regression analysis. Altman [5] established a default probability prediction model to predict the default probability of enterprises according to the ratio of retained earnings to total assets and other financial indicators. Douglas A et al. [6] found that net cash flows from operating activities, investing activities and financing activities are important sources for enterprises to repay debts. Sun [7] selected eight parts such as financial structure, growth level and profitability when identifying the financial distress of enterprises.

Beaver [8] argued that relying solely on financial factors to study bond default risk is inadequate due to challenges in obtaining accurate and reliable financial information. The research by Sadiq et al. [9] further suggests that the macroeconomic environment significantly influences corporate credit risk. In terms of the selection of non-financial indicators, Chava and Jarrow [10] found industry factors can affect the operation, profitability and development of enterprises in the same industry and industry cycle, and thus lead to bond default. Wu [11] selected three levels of non-financial indicators: macro, industry and enterprise. Eugene [12] and Wilson [13] discovered that when the macroeconomic environment is favorable, *per capita* disposable income grows more rapidly, which in turn leads to an increase in investment expenditure. The demand for bond investment rises, and enterprises have a stronger ability to repay principal and interest on bonds, with a lower risk of default. Lu Jun et al. [14] found that macroeconomic indicators such as GDP, CPI and M1 growth rate of narrow money supply had a significant

TABLE 3 Results of descriptive statistics.

Index	Mean	Std	Min	25%	50%	75%	Max
M1	518184.46	110831.42	327220.2	435628.1	543867.4	607266.8	674374.8
M2	1830854.03	445101.38	1,160,687	1,479,418	1,785,922	2,169,763	2,664,321
Business prosperity index	120.95	10.13	83.66	114.3	122.3	126.9	145.9
EPU	5.43	0.52	4.33	4.93	5.58	5.78	6.21
GDP	106.68	2.35	94.6	106.1	107	108	112.85
CPI	101.93	0.59	100.9	101.6	102	102.1	102.9
Local government debt	42.67	28.47	0	27	37	54	190
Proportion of shares held by major shareholders	64.63	30.42	4.03	38	64.84	100	100
Return on equity	3.29	33.99	−2204.89	1.2	4.25	8.54	183.86
EBIT/Total operating revenue	20.91	191.60	−4169.95	4.6	10.36	21.88	16185.02
Sales expenses/Total operating revenue	4.27	7.92	−2.27	0.79	2.35	4.96	336.7
Financial expenses/Total operating revenue	8.55	74.61	−1047.62	1.14	2.935	6.54	4725.95
Asset impairment loss/Total operating revenue	2.10	28.55	−431.47	−0.1	0.05	0.66	1534.54
Net income from operating activities/Total profit	−2.13	7186.05	−174767	23.1975	71.375	91.9525	650990.3
Net non-operating income and expenses/Total profit	49.67	1320.35	−4147.21	−0.13	1.15	8.425	112496.3
Income tax/Total profit	40.09	408.15	−2395.67	16.43	23.35	28.31	27516.8
Cash received from sales of goods and provision of services/Operating revenue	537.70	30087.72	−16.6	94.59	105.38	114.62	2,902,529
Long-term asset suitability rate	927857.00	89,754,481.00	−61399.8	133.2275	217.1	477.74	9.13E+09
Current assets/Total assets	47.90	23.62	0.61	29.2475	46.715	65.305	100
Current liabilities/Total liabilities	77.14	1108.30	0.62	51.83	67.86	81.24	109084.6
Quick ratio	1.08	7.68	0.02	0.55	0.79	1.1	707.8
Tangible assets/Interest-bearing debt	299.90	21570.23	−6.75	0.215	0.52	1.09	1,866,983
Net cash flows from operating activities/Non-current liabilities	−163.22	17768.32	−1268235	−0.06	7.62	26.61	219800.2
Tangible net worth debt ratio	−91.80	37889.14	−2971202	141.42	295.91	607.64	1,077,645
Inventory turnover rate	85.00	2158.91	−267.72	1.19	3.35	7.8625	136849.2
Accounts receivable turnover rate	100.31	2867.55	−1195.47	3.17	6.97	17.5875	219474.1
Accounts payable turnover rate	9435.04	675545.30	−10.08	1.97	3.81	7.985	58,719,863
Cash turnover rate	1053.23	12497.48	−5701.92	188.9	441.92	895.275	748911.4
Total operating revenue (year-on-year growth rate)	29.37	642.10	−100	−3.485	8.435	23.29	57909.39
Net profit (year-on-year growth rate)	−29.42	2996.55	−164908	−24.925	9.01	45.125	87598.59
Net assets (year-on-year growth rate)	9.30	115.05	−7951.27	1.38	7.33	16.51	1302.31
Total assets (year-on-year growth rate)	15.10	80.46	−2804.9	2.2175	9.45	19.52	4819.76

(Continued on the following page)

TABLE 3 (*Continued*) Results of descriptive statistics.

Index	Mean	Std	Min	25%	50%	75%	Max
Net cash flow (year-on-year growth rate)	523.87	19662.41	−112345	−122.9	22.01	165.65	1,554,680
Cash and cash equivalents/Short-term debt	33.33	1884.41	0	0.3	0.56	1.12	164211.5
Proportion of cash flows from financing activities	−1832.54	192496.20	−1.30E+07	−95.12	65.69	243.29	4,024,864
Proportion of cash flows from investing activities	1228.81	94828.06	−1701245	−170.4	−5.24	124.985	6,091,123
Proportion of cash flows from operating activities	703.52	124983.90	−4570310	−118.4	38.67	212.01	7,347,279

impact on companies falling into financial crisis. Chen [15] studies the relationship between macroeconomic environment and credit risk; At the macro level, LIU [16] selected CPI, social financing scale, market interest rate and so on.

Bonds belong to corporates' debt; debt repayment ability is closely related to the liquidity of enterprises. Money supply is a measure of market liquidity, it has an important impact on the price of various assets in the financial market, and its change will directly affect the economic environment. China's infrastructure projects account for a large share of the economy, money supply tends to outpace economic data. Fan et al. [17] studied the relationship between the excess rate of return of bonds and macroeconomic variables, and believed bond prices and market interest rates were mutually determined, and market interest rates were affected by factors such as money supply M1 and actual consumption. Through empirical evidence, they proved that changes in market interest rates might affect the expected excess rate of return of bonds. Based on the Merton structured model and the data of China from 2000 to 2010. Dai and Sun [18] studied the factors affecting the credit spread of corporate bonds in China's Shanghai and Shenzhen bond markets, and came to the conclusion that the issuance of M1 had a positive impact on the credit spread of corporate bonds. The fluctuation of stock index can indicate the current macroeconomic situation of economy and society to a certain extent, and the major changes of economic data will be reflected in the stock index. The stock market index is the vane of the stock price market, so the stock market index can also indicate the macro-economic situation.

In the selection of machine learning metrics, the causal relationship between metrics is a problem that should be given due attention. Some scholars have attempted to verify the relationship between the indicators and the predicted targets through causal relationship tests. In the field of transportation, Zhang et al. [19] studied the traffic speed of taxi in Changsha, China's urban road network, and tested the spatiotemporal causality between various links through nonlinear Granger causality, providing reliable guidance for MTL model to select information features from the whole network link. Luo [20] applied nonlinear Granger causality test to explore the causal relationship between traffic areas, built a multi-task deep learning model framework with long short-term memory (LSTM) as the core neural unit, and verified it with New York City taxi trip data. In the field of economics and finance, Wei et al. [21] selected the three variables with the highest correlation with the exchange rate among 16 macroeconomic variables including import, export and foreign exchange reserves through Granger causality test analysis, and used KELM to make a medium-term forecast of RMB against US dollar. The results show that the set learning method based on KELM outperforms all other benchmark models in different prediction periods. Xu et al. [22] combined Granger causality test with support vector machine and found that among the 15 stocks studied, the after-hours collective sentiment of nine stocks had a strong predictive effect on stock price changes the next day. In the field of medicine, Almalaq et al. [23] use Granger causality test to determine the activation region related to the verbal fluency task in human EEG, and train the classifier between subjects with Parkinson's disease and healthy control group in combination with support vector machine (SVM). In the field of physics, H et al. [24] introduced Granger causality to analyze the relationship between sensor variables, and selected variables with greater Granger causality relationship with sensor loss data as the input vector of the extreme learning machine. In the field of ecological environment, Vazquez-Patino et al. [25] explored the improvements in interpretability and robustness of models using causal selection predictors, comparing three cause-based methods with ML's four standard predictor selection methods using rainfall data from the Andean basin in Ecuador. Li et al. [26] screened the key environmental factors affecting sea ice concentration based on causal relationship (KGC), and used a variety of machine learning (ML) algorithms to make short-term sea ice prediction.

The literature reviewed highlights the related factors of bond risks and the application of causal analysis in machine learning. It is evident that scholars across various fields emphasize the importance of the correlation between selected indicators and the target when applying machine learning and deep learning algorithms. Through an analysis of the factors influencing corporate bond defaults, it has been found that key factors related to identifying corporate bond credit status include the financial metrics of the enterprise itself, which are the most representative indicators of its operational conditions. Non-financial indicators, such as macroeconomic factors, have also garnered attention from relevant scholars. Based on these findings, this paper employs causal analysis to select critical predictive indicators in the field of bond risk identification to validate the effectiveness of the chosen macroeconomic factors. Additionally, association mining is incorporated to describe the characteristics of risk aggregation. On this foundation, financial data and associated macroeconomic factors are selected, and ensemble learning algorithms are utilized to predict the credit status of

TABLE 4 Bond risk indicator system.

Factor	Indicator	Indicator	Indicator	Factor	Indicator
Profitability	Return on Equity (ROE)	Growth capacity	Total operating revenue (year-on-year growth rate)	Profit quality	Net income from operating activities/Total profit
	Ebit/Total operating revenue		Net assets (year-on-year growth rate)		Net non-operating income and expenses/Total profit
	Sales expenses/Total operating revenue		Net cash flow (year-on-year growth rate)		Income tax/Total profit
	Asset impairment loss/Total operating revenue		Net profit (year-on-year growth rate)	Local economy	GDP index
	Financial expenses/Total operating revenue		Total assets (year-on-year growth rate)		Local government debt
Cash flow	Net cash flow from operating activities ratio	Bond	Bond issuance rating	Corporate governance	Shareholding ratio of shareholders
	Net cash flow from financing activities ratio				Company type
	Net cash flow from investing activities ratio		Rating before maturity (default)	Macroeconomic	CPI, M1, M2, EPU
	Cash received from sales of goods Provision of services/Operating revenue		Rating change		
Capital structure	Tangible assets/Interest-bearing debt	Debt-paying ability	Quick ratio	Operating capacity	Inventory turnover rate
	Long-term asset suitability rate		Net cash flow from operating activities/non-current liabilities		Accounts receivable turnover rate
	Current assets/Total assets		Tangible net worth debt ratio		Cash turnover rate
	Current liabilities/Total liabilities		Cash And Cash Equivalents/Short-term debt		Accounts payable turnover rate
Industry impact	Business prosperity index				

specific bonds issued by enterprises, thereby enhancing the accuracy of identification. The main contributions of this paper lie in the following aspects: In terms of algorithm selection, traditional methods (such as logistic regression and probit models) perform poorly in handling nonlinear relationships and complex data structures. They have limited capacity to process high-dimensional data and are unable to capture the interactions among variables. We employed multiple ensemble machine learning algorithms and selected the optimal model through comparative analysis, significantly improving the prediction accuracy. In the construction of the indicator system, previous studies typically used only a single data source, lacking the integration of multi-source data. We developed a multi-dimensional bond default identification system that integrates macroeconomic indicators and market data, providing more comprehensive data support. Regarding the selection of macro indicators, previous studies often directly

TABLE 5 Stability test.

Variables	Original sequence	First difference sequence
Default	Non-stationary time series	stationary time series
M1	Non-stationary time series	stationary time series
M2	Non-stationary time series	stationary time series
HS300	Non-stationary time series	stationary time series
CPI	Non-stationary time series	stationary time series

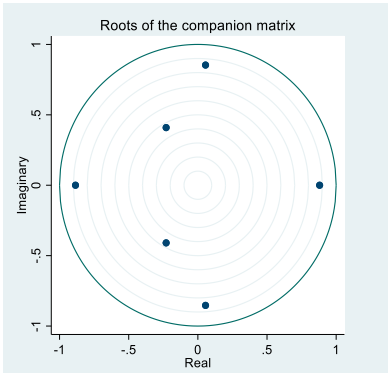


FIGURE 2
VAR stability test.

TABLE 6 ADF test results.

Variables	ADF	P-value	Critical value 1%
d.lnM1	−9.751	0	−3.689
d.lnM2	−6.169	0	−3.689
d.lnDefault	−8.374	0	−4.297

incorporated macro indicators into the prediction model without clarifying their relationship with bond defaults. This paper aims to demonstrate the causal relationship between macroeconomic indicators and bond defaults.

3 Research design and methodology

This article conducts bond risk identification. From a macro perspective, We use Granger causality test to analyze the causal relationship between macroeconomic indicators and the amount of bond default, and association rule mining is used to grasp the risk characteristics of bonds. From a micro perspective, individual bond samples are taken as the research objects. Combined with combinational learning classification algorithm, credit default problems of corporate bonds issued by companies in China are predicted. We select the data of corporate credit bonds issued by Chinese companies from 2014 to 2022. Including short-term financing bills, corporate bonds, enterprise bonds, medium-term notes, and asset-backed securities The main research objectives are divided into three parts: First, for the selected macroeconomic factors, we apply co-integration test and Granger test to verify the causal relationship between macroeconomic and bond default; Secondly, since most machine learning algorithms have limited interpretation of text type indicators, we use association rule mining algorithm to carry out visual cross-analysis of the characteristics of defaulted bonds, and grasp the distribution of default risk from a macro perspective. Finally, based on the verified results, we use the combinatorial learning model to forecast corporate bond default, compare the forecast results before and after adding macro factors, and conduct in-depth analysis of default factors. (Figure 1).

3.1 Causal analysis

VAR constructs the econometric analysis model by taking each endogenous variable in the economic system as a function of the other endogenous variables and their lagging values. It can deal with the estimation problems caused by endogenous variables as explained variables, identify and estimate the interdependence of endogenous variables or interactive spillover relationships. In Formula 1, y_t is a k-dimensional endogenous variable, x_t is an exogenous variable, p is the number of lag periods, A and B respectively represent matrices composed of estimated coefficients, and ε_t represents the error term. In the context of bond default analysis, VAR is used to analyze the relationship between Default and M1, M2, Default and CPI.

$$y_t = A_1 y_{t-1} + \ldots\ldots + A_\beta y_{t-p} + Bx_t + \varepsilon_t (t \in (1, T)) \quad (1)$$

3.2 Algorithm and evaluation index

APRIOR is used for association rule mining, APRIOR is a frequently used algorithm for mining frequent itemset, which was formally proposed by Agrawal et al [27]. In 1994. It features multiple scans of the database, a large scale of candidate items, and high computational cost for support calculation. By adopting an iterative approach, the algorithm steps are divided into two parts: "connection" and "pruning".

The reasons for choosing decision trees as the base learners are as follows: Compared with linear regression and other methods, decision trees have no strict requirements for data distribution, are robust to missing and noisy data, and can automatically handle and reduce the impact of outliers. Compared with neural network algorithms, decision tree models have an intuitive structure, are easy to interpret and visualize, provide a clear decision path for research problems, and have strong interpretability. Compared with algorithms such as KNN, they can efficiently handle high-dimensional data and reduce the influence of redundant features. Ensemble learning improves model performance by combining multiple base learners. Compared with a single machine learning model, ensemble learning integrates the prediction results of multiple models to reduce the bias and variance

TABLE 7 Trace test of Default, M1 and M2.

Maximum Rank	Params	LL	Eigenvalue	Trace statistic	Critical Value 5%
0	15	136.04948		76.7476	34.55
1	20	160.25203	0.79017	28.3425	18.17
2	23	170.15279	0.47205	8.5410	3.74
3	24	174.42327	0.24082		

TABLE 8 Default, M1, M2 maximum characteristic root test.

Maximum rank	Params	LL	------Eigenvalue----- Maximum		Critical value 5%
0	15	136.04948		48.4051	23.78
1	20	160.25203	0.79017	19.8015	16.87
2	23	170.15279	0.47205	8.5410	3.74
3	24	174.42327	0.24082		

TABLE 9 D.lnDefault D.lnM1 D.lnM2 Optimal lag order.

Lag	LL	LR	DF	P	FPE	AIC	HQIC	SBIC
0	123.521	1	1	1	3.7e-08	−8.60862	−8.56498	−8.46588
1	135.894	24.747	9	0.003	2.9e-08	−8.84957	−8.67503	−8.27862
2	162.614	53.441	9	0.000	8.4e-09*	−10.1153*	−9.80987*	−9.11617*
3	167.042	8.855	9	0.451	1.2e-08	−9.78871	−9.35235	−8.36135
4	179.946	25.808*	9	0.002	1.1e-08	−10.0676	−9.5003	−8.212

*Optimal lag.

of a single model, thereby improving overall prediction accuracy. By majority voting or weighted averaging, it reduces the influence of data noise, thus having stronger robustness to noisy data and outliers. A single model may overfit the training data, while ensemble learning achieves a balance between model complexity and generalization ability by integrating the prediction results of multiple models. The research purpose of this paper is to identify the key factors of bond default and find an algorithm that can accurately predict bond default in practice. Given the robustness, interpretability, efficiency, and flexibility of decision tree ensemble algorithms, we choose the decision tree-based ensemble learning algorithm to predict bond risk.

Bagging and Boosting combined learning algorithms are used. Bagging is a homogeneous estimator composed of many decision trees. Boosting base learner construction has a sequence [28]. Bond risk prediction employs decision trees, the bagging algorithm based on decision trees - random forest, and boosting algorithms based on decision trees, including Xgboost, Catboost, GBDT and Hist-GBDT. GBDT (Gradient Boosting Decision Tree), also known as MART (Multiple Additive Regression Tree), The concept consists of Regression Decision Tree (DT), Gradient Boosting (GB) and Shrinkage. GBDT constructs a set of weak learners that add up the results of multiple decision trees as the final predictive output. GBDT iterates the model through gradient descent to reduce the impact of the cost function. By calculating the negative gradient of the loss function to construct residuals, it gradually reduces the value of the loss function, thereby obtaining increasingly accurate learners. GBDT can be expressed in the Equation 2. Xgboost consists of two parts: empirical risk and structural risk (regularization term), and it uses the forward stagewise algorithm to gradually optimize the classifier. Xgboost can be expressed as Equation 3. By effectively processing categorical features and introducing ranking boosting strategies, CatBoost addresses gradient bias and prediction shift issues [29]. CatBoost can efficiently handle class quantities, prevent overfitting, and train models with high accuracy. Its built-in algorithm can automatically transform categorical features into numerical features, and combine features according to the intrinsic relationship of features

TABLE 10 M1, M2 and Default Granger causality test.

Equation	Excluded	chi2	Df	Prob > chi2
D_lnDefault	D.lnM1	5.0613	2	0.08
D_lnDefault	D.lnM2	0.2828	2	0.868
D_lnDefault	ALL	6.3535	4	0.174
D_lnM1	D.lnDefault	6.8288	2	0.033
D_lnM1	D.lnM2	2.9831	2	0.225
D_lnM1	ALL	12.35	4	0.015
D_lnM2	D.lnDefault	5.3539	2	0.069
D_lnM2	D.lnM1	29.059	2	0
D_lnM2	ALL	34.044	4	0

to enrich the feature dimension. Its biggest characteristic lies in its ranking idea. HistGradientBoostingClassifier uses histogram data structure to arrange data samples implicitly, and only the largest split nodes are considered in the tree building process, so the number of split nodes is small, and only the initial input data needs to be sorted. Most parts of the HistGradientBoostingClassifier algorithm are implemented in parallel, which can effectively improve the construction efficiency of base classifier.

$$F(x) = F_M(x) + \sum_{m=1}^{M} \eta \sum_{j=1}^{J} \gamma_{jm} \Pi\left(x \in R_{jm}\right) \tag{2}$$

$$Obj = \sum_{i=1}^{N} L|F_m(x_i), y_i| + \sum_{j=1}^{N} \Omega\left(f_j\right) = \sum_{i=1}^{N} L|F_{m-1}(x_i) + f_m(x_i), y_i| \tag{3}$$

Common evaluation metrics in machine learning include confusion matrices and AUC curves. Bond risk assessment is a binary classification problem, so a confusion matrix can be used to represent the four possible outcomes of model operation, namely, FP, TP, FN, and TN, as shown in Table 1. Through the confusion matrix, accuracy, precision, recall, and F-score can be calculated.

$$Accurary = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$\text{Pr}ecision = \frac{TP}{TP + FP} \tag{5}$$

$$\text{Re}call = \frac{TP}{TP + FN} \tag{6}$$

$$FPR = \frac{FP}{FP + TN} \tag{7}$$

$$F_{Score} = \frac{2\ \text{Pr}ecision * \text{Re}call}{\text{Pr}ecision + \text{Re}call} \tag{8}$$

Accuracy is the ratio of the number of correctly predicted samples to the total number of samples (Formula 4). Precision is the ratio of the number of samples predicted as positive and actually positive to the total number of samples predicted as positive (Formula 5). Recall is the ratio of the number of samples that are truly positive and correctly predicted as positive to the total number of samples that are actually positive (Formula 6). False positive rate is the ratio of the number of samples predicted as positive but actually negative to the total number of samples that are actually negative (Formula 7). The F1 score is the harmonic mean of precision and recall (Formula 8). The horizontal axis of the ROC curve is the false positive rate, and the vertical axis is the recall rate. The area under the ROC curve is denoted as AUC. Generally, when AUC >0.5, the model results are considered meaningful. Recall and AUC are used to evaluate the accuracy of classification models. Before conducting machine learning experiments, parameter optimization is required. The goal of parameter optimization is to improve the prediction accuracy by optimizing the model's hyperparameters, prevent overfitting, and enhance the model's generalization ability. We use grid search and cross-validation in Python's scikit-learn, with AUC as the evaluation metric, to determine the reasonable values of hyperparameters. By traversing the predefined parameters, the parameters with the best performance are selected. The detailed parameter adjustment results are shown in Table 2.

4 Data and analysis

The research data of this article is sourced from Wind database and the National Bureau of Statistics of China. Table 3 is the descriptive statistics of the data. Through the analysis of literature review and the sorting out of the evolution process of bond risks, the factors selected in this article include macroeconomics, bond characteristics, local economy, industry characteristics, and the financial status of bond issuers. When analyzing the financial bond market, macroeconomic indicators and market indices are important influencing factors. M1 and M2 are key indicators for measuring the money supply. The changes in M1 and M2 directly affect market liquidity and interest rates, thereby having a significant impact on the bond market. CPI is an important indicator for measuring the level of inflation. Inflation expectations directly affect bond yields and prices. Based on the previous studies by Lu et al., Fan et al., and Dai and Sun, monetary indicators and CPI can affect bond defaults, so we incorporate them into the macroeconomic dimension. The economic environment of the region where the bond issuer is located is also an important influencing factor, so we include local GDP in the local economic dimension. At the same time, the characteristics of the bond itself cannot be ignored. The financial status of the bond issuer includes six aspects: capital structure, profitability, cash flow, debt-paying ability, operational ability, and development ability. There is an inherent correlation and synchronous change among financial indicators, and the contribution of financial indicators will affect the stability, explanatory power, and predictive performance of the model. Therefore, we conducted a correlation test and deleted the correlated indicators with a correlation coefficient greater than 0.6. We used the indicators that passed the correlation test to construct a bond risk identification index system (Table 4). Based on the principles of systematisms, scientific, and feasibility of the indicators, after data processing, a total of 10,516 bond data were obtained, including 657 default samples. The training set and test set were divided in a 7:3 ratio.

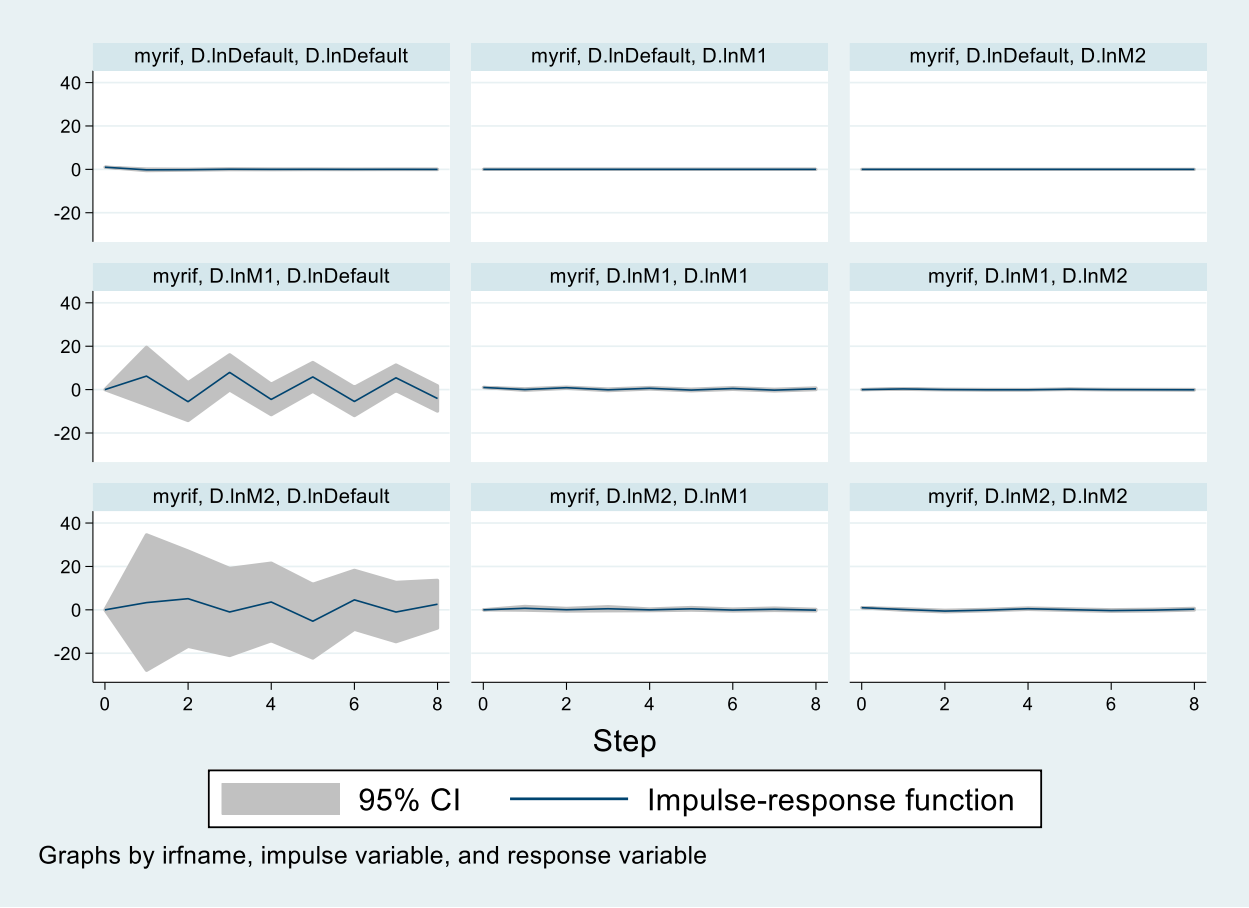


FIGURE 3
Default and M1 and M2 pulse response.

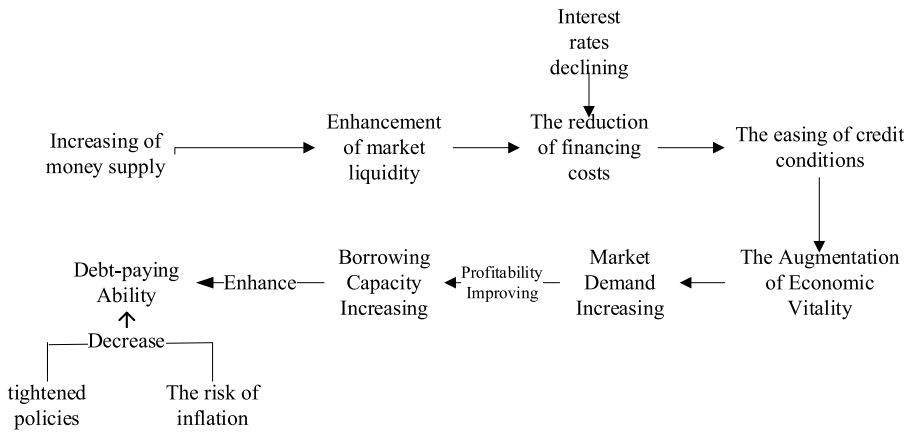


FIGURE 4
Macro risk transmission map.

4.1 Granger causality test

The purpose of the stability test is to assess whether these time series data are stationary. Stationarity tests were conducted on the default rate, M1, M2, and CPI for each quarter from 2014 to 2022. After first-order differencing, the time series of all variables became stable (Table 5). We select the reciprocal distribution of feature roots to verify the stationarity of VAR. When the reciprocal of

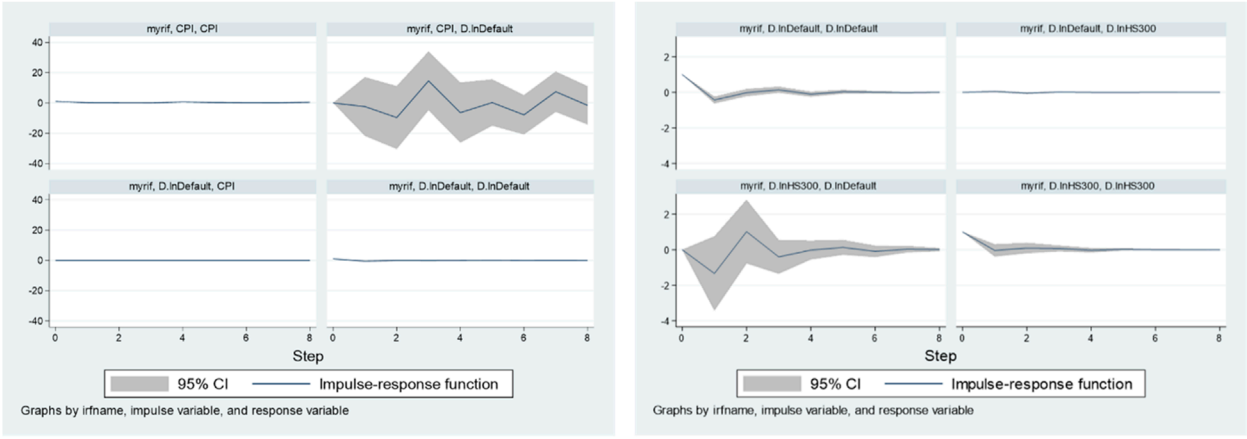


FIGURE 5
Default and CPI, Default and HS300 pulse response.

TABLE 11 Riskless bond association rules.

Support	{Local state-owned enterprises} → {unlisted companies, Riskless bonds}
Confidence	{Unsecured, Central state-owned enterprises} → {North China, Riskless bonds}
Lift	{Central State-owned enterprises} → {North China, Riskless bonds}
Frequency	Guarantee > Company Nature > Region > Industry

TABLE 12 Risk bond association rules.

Support	{Unlisted companies} → {risky bonds, unsecured} {Unsecured} → {risky bonds, unlisted companies}
Confidence	{unlisted companies, materials} → {risky bonds, unsecured}
Lift	{East China} → {Risky bonds, unsecured, unlisted companies, private enterprises}
Frequency	No guarantee, listed or not and regions are High frequency attributes

feature roots is distributed in the unit circle, the model is stable. The results show that all the reciprocal of the characteristic roots are within the unit circle (Figure 2), indicating that the model is stable. Additionally, we conducted the ADF test (Table 6), and the ADF statistics of all variables are significant, suggesting that the data series is stable. Therefore, the stability of the VAR has been fully verified. The VAR system meets the conditions for impulse response analysis.

The cointegration test of Default, M1 and M2 is carried out. The trace test results and the maximum eigen Root results show that the cointegration relationship between variables is significant (Table 7; Table 8). Table 9 shows the selection results of the optimal lag order. LR = 4, FPE, AIC, HQIC, SBIC = 2. According to the characteristics of each index, the lag order is determined as 2. The data is quarterly, which takes into account the causal relationship between factors with a lag of 6 months and defaults. According to the results of cointegration test, Granger causality test is used to explore the causal relationship between the above variables. The results are shown in Table 10. M1, M2 have an impact on bond default respectively, while M1 and M2 have an impact on the overall system respectively. Default only has an impact on M1. Therefore, there is a two-way causal relationship between M1 and the default amount of bonds, and a one-way causal relationship between M2 and the default amount. M2 has an effect on M1, while M1 has no effect on M2.

The impulse response graph is shown in Figure 3. The horizontal axis of the impulse response function represents the lag period, and the vertical axis reflects the response fluctuation level of the target variable to the unit shock. The solid line and the shaded area represent the impulse response function. Based on the corresponding impulse images, we analyzed the response process of M1 and M2 to default. The impact of M1 on default generally shows periodic fluctuations, with positive and negative effects being symmetrical. When M1 changes, default first responds positively and then moves in the opposite direction, with a period of approximately 2, that is, 1 year. After moderate symmetrical fluctuations, it tends to converge. Compared with M1, the response path of M2 to the default amount has smaller fluctuations. When the impact of M2 on default does not exceed 1 year, default first rises slowly in a positive direction and then drops in a negative direction. After that, the fluctuation period of the impact gradually shortens. Based on the above findings, we demonstrated the connection between macroeconomic monetary indicators and bond risk and sorted out the risk evolution process (Figure 4). Loose monetary policy leads to an increase in the money supply, thereby enhancing market liquidity, reducing financing costs, easing credit conditions, and gradually affecting the short-term and long-term debt repayment capabilities of enterprises; the strengthening of economic activities, the rise in asset prices, and the improvement in consumer demand will further enhance the debt repayment capabilities of enterprises. We also verified the monetary supply and credit transmission mechanism through impulse response analysis.

TABLE 13 Risk bonds strongly associated attributes.

Dimension	Value	Characteristics
Enterprise Dimension	Private enterprise	The confidence of 7 rules is higher than 95%
	listed or not	The top five rules are related to non-listed companies
Industry	Industrial sector	Industrial bonds account for a large proportion
	Material industry	Confidence of {materials}→{risky bonds, unsecured} = 97.47%
	Optional consumption	{Optional consumer sector}→{risky bonds, unsecured}
	Real estate	{Real estate}→{risky bonds, private enterprises}
Region	East China	Developed economy; Many private enterprises
	South China	
Guarantor or not	No Guarantor + Industrial	Often associated with others in association rules
	East China + Industrial	

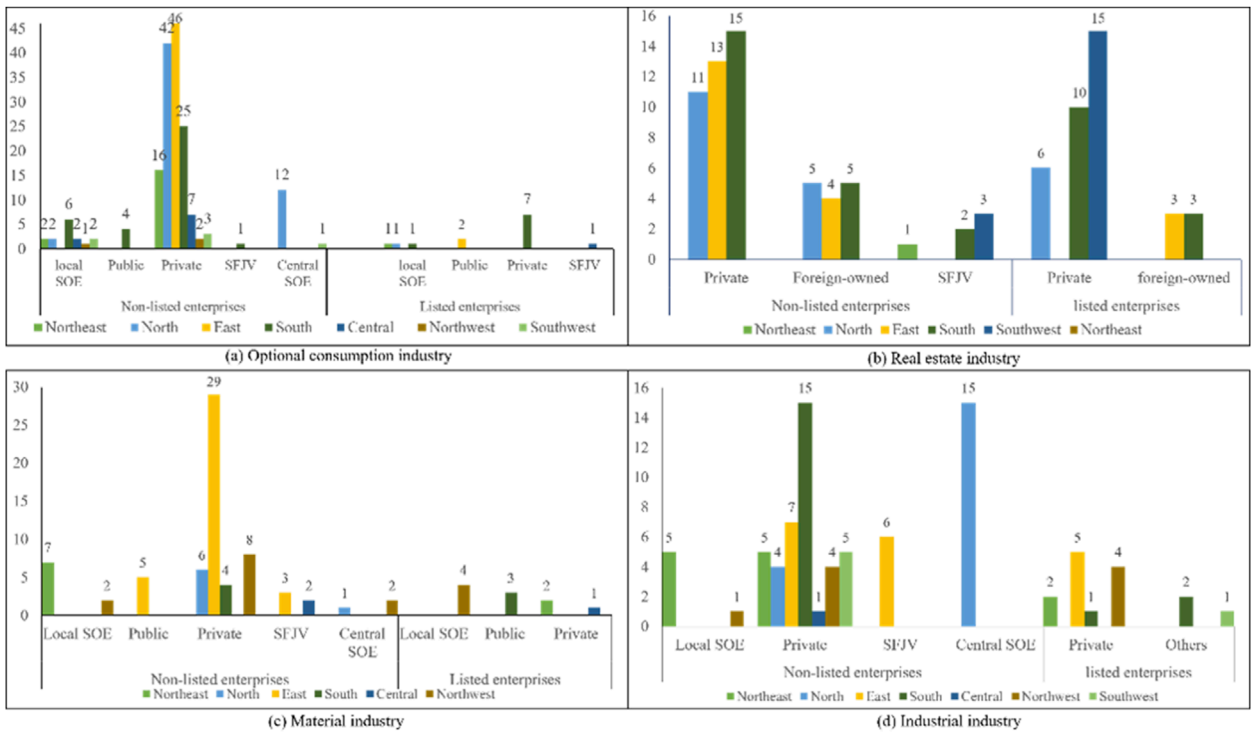


FIGURE 6
Risk distribution of default industry. **(a)** Distribution of characteristics of defaulted bonds in optional consumption industry; **(b)** Distribution of characteristics of defaulted bonds in real estate industry; **(c)** Distribution of characteristics of defaulted bonds in Material industry; **(d)** Distribution of characteristics of defaulted bonds in Industrial industry. SFJV, Sino-foreign joint venture.

Following the above steps, we explore the impact of stock market factors on bond market and the impact of CPI on bond default. The pulse response plots and results are shown in Figure 5. The impact of CPI on bond default lasts for a long time. From the overall trend, the pulse effect given by CPI firstly have a negative impact on default, and then tend to be positive. HS300 index is selected as the characteristic index of stock market. There is Granger causality between HS300 and default. Under the influence of HS300, default first shows a relatively rapid negative change and then changes into a positive change, with a cycle of about 3 quarters. After a positive and negative influence appears alternately, the influence on the default amount tends to be stable.

TABLE 14 Classification prediction results of bond risk models.

	Model		Decision	RandomForest	GBDT	Hist-GBDT	Catboost
SOEs	Accuracy	Matured bonds	0.99	0.99	1	1	1
		Risky bonds	0.50	0.82	0.49	0.85	0.84
	Recall rate	Matured bonds	0.99	1	0.99	1	1
		Risky bonds	0.65	0.68	0.85	0.85	0.94
	F1-score	Matured bonds	0.99	1	0.99	1	1
		Risky bonds	0.56	0.74	0.62	0.85	0.89
	Accuracy rate		98.36%	99.23%	98.31%	99.52%	99.61%
Non-SOEs	Accuracy	Matured bonds	0.97	0.98	0.99	0.99	0.99
		Risky bonds	0.81	0.93	0.80	0.92	0.95
	Recall rate	Matured bonds	0.96	0.99	0.96	0.99	0.99
		Risky bonds	0.86	0.88	0.95	0.95	0.95
	F1-score	Matured bonds	0.97	0.98	0.97	0.99	0.99
		Risky bonds	0.83	0.90	0.87	0.94	0.95
	Accuracy rate		94.26%	97.04%	95.47%	97.96%	98.33%

4.2 Characteristic analysis of defaulted bonds

Based on credit status, corporate bonds can be classified into matured bonds and defaulted bonds. By comparing the five attributes of the issuer's location, company characteristics, primary industry, whether there is a guarantor, and whether the issuer is a listed company, we analyzed the characteristics of defaulted bonds and obtained the association rules with the highest support, confidence, and lift values, as well as the attribute frequency ranking, between matured bonds and defaulted bonds. The association rules for risky bonds are shown in Table 11, and those for risk-free bonds are presented in Table 12. Based on the risk bond association rule, we observe that the credit risk of corporate bonds is more concentrated in four key industries: real estate, materials, discretionary consumption, and industrial sectors (Table 13). Utilizing association rules, we conduct a cross-statistical analysis of company attributes and issuer locations across these four industries. Overall, non-listed issuers exhibit higher risk levels compared to their listed counterparts in each industry.

The results of the cross-tabulation analysis of risk bonds with listing status and geographic distribution attributes for the industrial, real estate, consumer discretionary and materials sectors are shown in Figure 6. The risk of listed companies in the optional consumption sector is lower than that of unlisted enterprises, and unlisted private enterprises in South China and central state-owned enterprises in North China are particularly worthy of attention. The risks of corporate bonds in the real estate industry are primarily concentrated within private enterprises. Whether the issuing entity is listed or not has a negligible impact on risk discrimination. The bonds issued by non-listed private enterprises in North China, East China, and South China, as well as those issued by listed private enterprises in North China, South China, and Southwest China, entail relatively significant risks. Nevertheless, the bonds issued by non-listed Sino-foreign joint venture real estate enterprises have lower risks than those issued by wholly foreign-owned enterprises in the real estate sector. The bonds issued by non-listed Sino-foreign joint venture real estate enterprises in Northeast China encounter the least risk exposure.

The risk associated with bonds issued by the industrial and materials sectors is predominantly concentrated in unlisted enterprises. Within the industrial sector, unlisted private enterprises in East China, North China, and South China exhibit higher risk levels, bond defaults among centrally-owned state enterprises in North China warrant attention. In contrast, bonds issued by locally-owned state enterprises and foreign joint ventures tend to be less risky. The risk profile of bonds issued by industrial enterprises demonstrates a strong regional clustering effect. In the material industry, the number of bond defaults of private enterprises in East China is the highest, because the material industry is greatly affected by the energy policy, which leads to the cash flow turnover problem of bond issuers in the material industry, especially private material enterprises. The corporate bond risk shows no substantial variation based on regional or other characteristics.

The above analysis has examined the distribution characteristics of defaulted bonds from an industry perspective. It can be concluded that industries with a higher correlation to bond defaults mainly have strong cyclical nature, high leverage, policy sensitivity, and external environmental uncertainty. During economic downturns or deteriorating market conditions, the profitability, cash flow, and financing capabilities of these industries are more vulnerable to

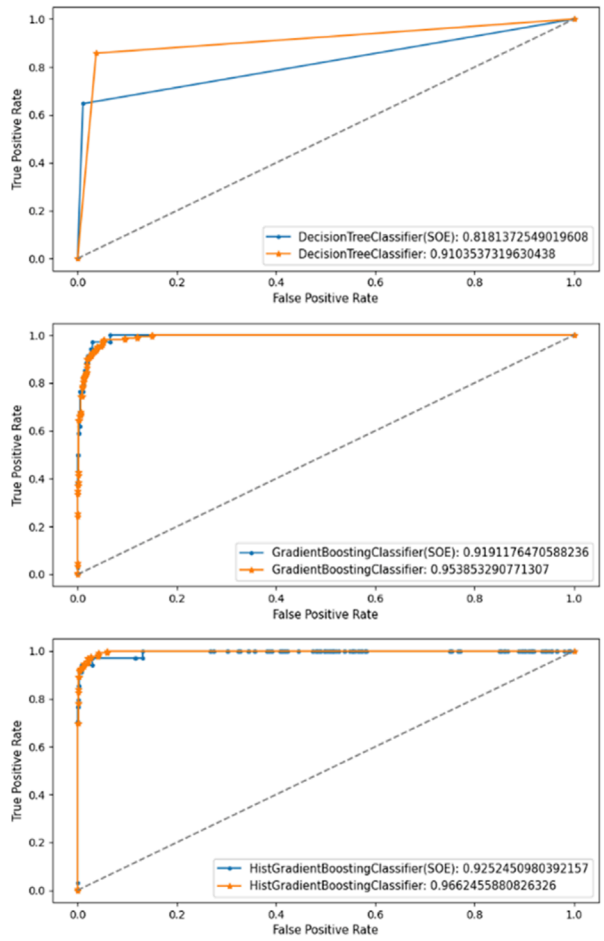


FIGURE 7
Comparison of ROC curves.

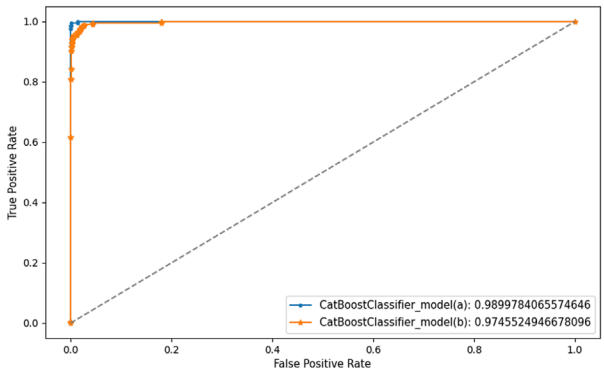


FIGURE 8
ROC curve for the ablation experiment.

shocks, thereby increasing the risk of default. From the perspective of whether they are listed or not, the average risk of bond issuance by listed companies is lower than that of unlisted companies. In terms of regional distribution, the characteristics of default regions are mostly concentrated in the eastern, northern, and southern parts of China, mainly due to the greater economic vitality and a larger number of enterprises in these regions.

The results of the above descriptive statistics are basically consistent with the analysis and association rule mining results, it confirms the effectiveness of association rule mining. The cross-analysis of the four industries can grasp the distribution and concentration of bond credit risk from a macro perspective, and assist investors to make relatively safe decisions when investing in public offering corporate bonds.

4.3 Bond default forecast

The attribute of enterprise frequently appears in the mining of association rules, and the results of association rules show that the bond risks of different industries vary in this dimension. Using machine learning models, we segmented the dataset into two groups based on the nature of the bond issuers: state-owned enterprises (SOEs) and non-state-owned enterprises (NSOEs). Bond risk prediction was then conducted on these two datasets, yielding distinct prediction outcomes. The results are shown in Table 14.

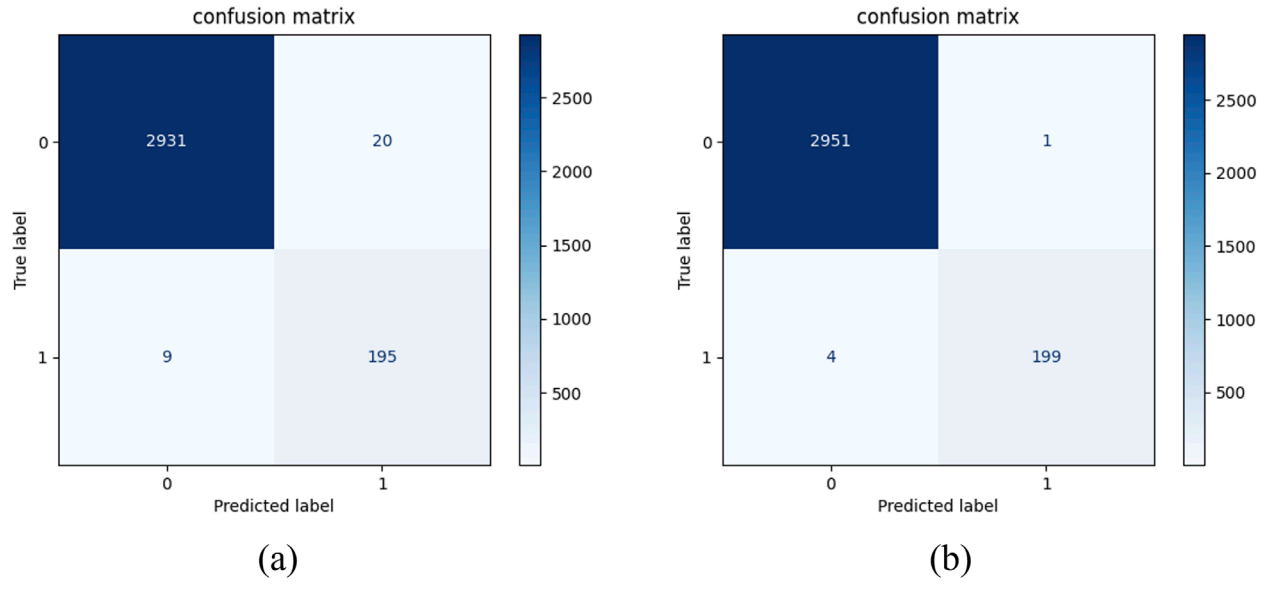


FIGURE 9
Comparison chart of confusion matrices for the ablation experiment. **(a)** Confusion matrix of Catboost before adding macro features; **(b)** Confusion matrix of Catboost after adding macro features.

By comparing the experimental outcomes of the two groups of data, it is observed that the accuracies of decision trees and GBDT in the two groups of experiments are not high. From the perspectives of precision, recall rate, and F1-score metrics, the prediction accuracies of random forests, HistGradientBoostingClassifier, and Catboost for non-state-owned bonds are higher than those for state-owned bonds. CatBoost performs best on the state bond dataset, while HistGradientBoostingClassifier and CatBoost are the top models for another group. According to the ROC curve analysis (Figure 7), the AUC values predicted by the decision tree and GBDT algorithm models for state-owned bonds are higher than those for non-state-owned enterprise bonds. Conversely, other models predict higher AUC values for non-state-owned bonds compared to state-owned enterprise bonds. Among the models, Random Forest and CatBoost exhibit similar performance across both datasets. The HistGradientBoostingClassifier achieves an AUC score that is 0.0916 higher on the non-state bond dataset compared to the state bond dataset. CatBoost performs best on the state bond dataset, while HistGradientBoostingClassifier outperforms other models on the non-state bond dataset.

In addition, to verify the conclusion of the var experiment, we selected the Catboost model and conducted ablation experiments on the indicators M1 and M2. The experimental results are shown in Figures 8, 9. The addition of the monetary indicator increases the AUC area of Catboost and reduces the number of bond misclassifications. This proves that m1, m2 has an enhancing effect on recognizing bond risk.

5 Results and discussion

The three experiments, VAR analysis, association rule mining, and descriptive cross-statistical analysis, analyzed the risks of credit bonds from the overall perspective of the bond market. Granger causality tests were conducted on the monetary indicators M1 and M2 and bond defaults, and the results confirmed the existence of a causal relationship between macroeconomic indicators and bond defaults. M1 and M2 have an impact on bond defaults: In light of the actual situation of China's bond market, China's interest rate liberalization is nearly complete, and the bond market has become an important platform for monetary policy operations. The interest rate transmission from the money market to the bond market is an important channel for monetary policy. M1 and M2 influence the bond market through liquidity effects and economic expectations, and their changes have an impact on the trend of the bond market. Bond defaults have an impact on M1: Against the backdrop of a high incidence of credit default events, defaulted bonds will lose liquidity due to trading suspension. Investors' risk aversion will lead them to choose to sell liquid bonds in their asset portfolios, thereby weakening the liquidity of high-quality bonds. Bond market risks will also lead to a decline in systemic risk preferences in the financial market, further affecting monetary policy.

Association rule mining and cross-analysis identified the current existence and concentration points of risks. Based on the results of association rule mining, we have identified four industries with concentrated bond risks: real estate, materials, consumer discretionary, and industrial. These four industries with high-risk correlation are highly consistent with the risk concentration points in China's bond market. The industrial and materials sectors have been affected by policies such as China's industrial structure upgrade and environmental protection production restrictions. As a result, some enterprises with overcapacity and high energy consumption and pollution have faced increased operational pressure, rising financing difficulties, greater capital turnover pressure, and continuous decline in profit levels, leading to accelerated bond risk release. The consumer discretionary sector is highly sensitive to economic cycles and is

TABLE 15 Feature importance ranking of catboost.

Feature	Importance ranking of state-owned bonds	Feature	Ranking of non-state-owned bonds
Net cash flows from operating activities/Non-current liabilities	33.40934	Net cash flows from operating activities/Non-current liabilities	40.15558
Net assets (year-on-year growth rate)	8.340914	Impairment loss on assets/Total operating revenue	5.10537
pca2	4.570496	Net non-operating income and expenses/Total profit	3.173927
Financial expenses/Total operating revenue	3.597293	Financial expenses/Total operating revenue	2.932994
Proportion of shares held by major shareholders	3.51573	Long-term asset suitability rate	2.766285
pca4	3.420728	Net assets (year-on-year growth rate) M1	2.420678
Net cash flows from operating activities as a percentage	3.065502	M1	2.365089
Asset impairment loss/Total operating revenue	2.900967	Total assets (year-on-year growth rate)	2.346182
pca5	2.535832	GDP	2.307728
Inventory turnover rate	2.440826	EPU	1.959694
pca1	2.311772	Current liabilities/Total liabilities	1.914631
GDP	2.085854	M2	1.774024
Current assets/Total assets pca7	1.986099	The shareholding ratio of major shareholders	1.758257
pca7	1.985943	pca2	1.746224
Accounts payable turnover ratio	1.789056	Local government debt	1.605419
Net cash flows from investing activities as a percentage of total cash flows	1.551838	Accounts payable turnover ratio	1.604
Net cash flows from financing activities as a percentage of total cash flows	1.477189	Tangible assets/interest-bearing debt	1.53493
Cash turnover ratio	1.398997	Proportion of net cash flows generated from operating activities	1.51413
Current liabilities/Total liabilities	1.364536	Proportion of net cash flows generated from investing activities	1.367484
Cash received from selling goods and providing services/Operating income	1.138044	Current assets/total assets	1.346553

mainly affected by the sluggish macroeconomy and weak consumer demand. For the real estate sector, it is because China's real estate policy regulation has become stricter in recent years. Regulations such as "houses are for living in, not for speculation" have suppressed housing price increases while also restricting the financing of real estate enterprises, leading to accelerated exposure of bond credit risks. According to cross-analysis, the risks of corporate bonds in the real estate industry are mainly concentrated in private enterprises, which confirms the current situation of two-tiered differentiation in the operation of Chinese real estate enterprises. Currently, the net bond financing of state-owned real estate enterprises in China has significantly increased, but the financing situation of some private real estate enterprises is not optimistic. At the same time, we have found that in each industry, the risk level of non-listed issuers is higher than that of listed companies. This is because listed companies have diverse financing channels, are subject to more supervision, and can

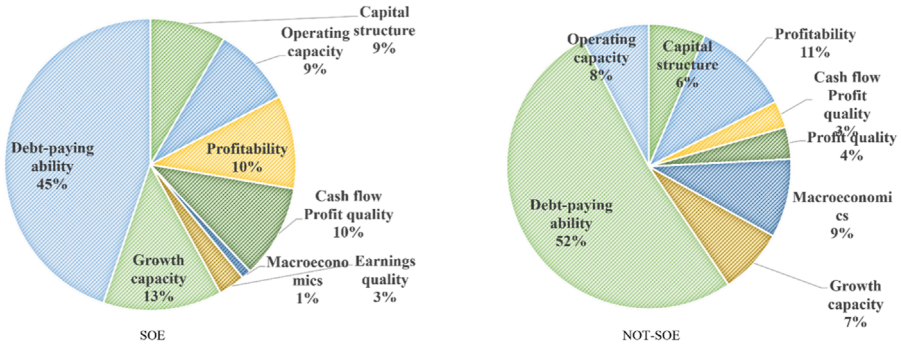


FIGURE 10
Comparison of risk factors.

disclose more information. Therefore, the credit bond risks of listed enterprises are usually smaller than those of non-listed enterprises. From a regional perspective, the economic development levels in South China and East China are relatively high, with a large number of private enterprises and concentrated bond issuers. Most enterprises are mainly engaged in trade, manufacturing, and export processing, and are more vulnerable to economic environment changes, thus leading to relatively concentrated default risks. These discussions provide valuable insights for subsequent focused attention and targeted adjustments in industries and regions experiencing frequent defaults.

With individual credit bonds as the research object and based on the aforementioned results of machine learning, we explore the influencing factors of state-owned and non-state-owned bonds, and further compare them for difference analysis. Since the AUC of Catboost is greater than 90% on both sets, we give the ranking of the top 20 feature importance of Catboost after removing the principal component index on the two data sets (Table 15). Through the feature importance weights, a comparison chart of risk factor weights for the two groups of experiments was obtained (Figure 10). Debt-paying ability can directly reflect the risk of credit bonds, and its proportion in private enterprises exceeds 50%. This indicates that financial indicators related to debt-paying ability most directly reflect bond risk. Capital structure also holds certain significance for the risk identification of credit bonds. This indicates that enterprises should pay attention to the investment and financing structure as well as the proportion of enterprise expenditures during the investment and financing process. A reasonable investment and financing ratio can maintain the stability of cash flow to diversify risks. For listed companies, the risk of bonds is smaller than that of issuing stocks, and it also provides them with a tool to diversify risks. Profitability reflects short-term solvency, while development capacity is demonstrated through indicators such as revenue growth rate, asset growth rate, and ROE, which also reflect long-term solvency. The combination of the two comprehensively reflects the solvency of an enterprise and represents its development potential, making them relatively important risk identification indicators. The proportion of macroeconomics in the two sets of data varies greatly. Non-private enterprises are greatly affected by macroeconomics and macro factors on their own operations. Private enterprises are more susceptible to policy adjustments. When the credit environment is relatively tight and the overall financing environment tightens, the liquidity pressure on private enterprises' funds will increase, which will significantly increase the risk of bond default. Therefore, issuers should make accurate predictions of the macro environment and the development positioning of their enterprises before issuing credit bonds.

6 Conclusion

This paper analyses bond default from the perspective of macro and micro combination. Based on the macro perspective, the logical transmission between macro factors and bond default is confirmed by causation, and the overall characteristics of bond default are grasped by the statistical research method combining association rule mining and descriptive statistics. Bonds are divided into no-risk bond group and risk bond group, and association rules are mined in four dimensions, corporate bond region, whether the issuing entity is listed, the issuing enterprise attribute, and whether there is a guarantee for corporate bonds, and rules with high correlation degree are obtained, the cross-analysis of risk factors was carried out in these four different dimensions. Microscopically, the bonds issued by enterprises were taken as the research object. The characteristics of defaulted bonds were analysed from the perspective of whether the issuer is a state-owned enterprise or not, and the key factors for bond risk identification and the integrated algorithm with the highest AUC were obtained. The conclusions are as follows:

(1) Through VAR and Granger causality test, we prove that bond default is correlated with M1 and M2;
(2) In view of the text attributes of bonds and bond issuers, correlation mining is conducted to find out the risk attributes strongly related to bond, and further conducted descriptive statistics on risk attributes to further grasp the risk characteristics and trends on a macro level.
(3) Bonds are grouped based on the risk correlation features obtained from association rules mining, and the credit status of bonds is predicted by combinational learning algorithm. The importance of risk attributes is ranked from the micro level according to the prediction results, and the validity of Granger causality test is further confirmed. At the same time,the combination learning model used in this paper has a high accuracy for bond default prediction, the accuracy rate of all

algorithms exceeds 90%. In terms of recall rate, the ensemble learning algorithm is better than the decision tree. The AUC of Catboost on both sample sets is greater than 90%. Therefore, we chose CatBoost to conduct a comparative experiment on the risk identification of bonds issued by state-owned and non-state-owned enterprises. By analysing the importance features of two groups of machine learning, the ranking of feature importance was obtained. The top-ranked feature was the solvency of enterprises, including important indicators such as "net cash flow from operating activities/non-current liabilities", indicating that the cash flow status of enterprises is extremely crucial for assessing the risk of corporate bonds. Macro factors are also key factors in bond risk identification. Non-state-owned enterprises are more affected by macro and their own operating conditions. In addition, in the ranking of the importance of default factors for non-state-owned bonds, both M1 and M2 are relatively important features. This results and the ablation experiments in machine learning indicate the correlation between bond default and M1 and M2, verifying the correctness of the aforementioned causal test from a micro perspective. Our research can provide a tool reference and guidance for risk management in the actual bond market.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YWZ: Data curation, Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. WC: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Merton RC. On the pricing of corporate debt: the risk structure of interest rates. *The J Finance* (1974) 29(2):449–70. doi:10.1111/j.1540-6261.1974.tb03058.x

2. Kedia S, Rajgopal S, Zhou X. Did going public impair Moody' s credit ratings? *J Financial Econ* (2014) 114(2):293–315. doi:10.1016/j.jfineco.2014.07.005

3. Yang B, Bin K, Li B, Julia Yu Y, Zhang J. Detecting accounting fraud in publicly traded US firms using a machine learning approach. *J Account Res* (2020) 58(1):199–235. doi:10.1111/1475-679X.12292

4. Beaver WH. Financial ratios as predictors of failure. *J Account Res* (1966) 4(12):71–111. doi:10.2307/2490171

5. Altman E, Fargher N, Kalotay E. A simple empirical model of equity- implied probabilities of default. *J Fixed Income* (2011) 20(3):71–85. doi:10.3905/jfi.2011.20.3.071

6. Douglas A, Huang AG, Vetzal KR. Cash flow volatility and corporate bond yield spreads. *Social Sci Electron Publishing* (2016) 46(2):417–58. doi:10.1007/s11156-014-0474-0

7. Sun J, Hui L, Fujita H, Fu B, Ai W. Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. *Inf Fusion* (2020)(54) 128–44. doi:10.1016/j.inffus.2019.07.006

8. Beaver WH, Mcnichols MF, Rhie JW. Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Rev Account Stud* (2005) 10(1):93–122. doi:10.1007/s11142-004-6341-9

9. Sadiq M, Alajlani S, Hussain MS, Ahmad R, Bashir F, Chupradit S. Impact of credit, liquidity, and systematic risk on financial structure: comparative investigation from sustainable production. *Environ Sci Pollut Res* (2022) 29(14):20963–75. doi:10.1007/s11356-021-17276-x

10. Chava S, Jarrow RA. Bankruptcy prediction with industry effects. *Eur Finance Rev* (2004) 8(4):537–69. doi:10.1007/s10679-004-6279-6

11. Yuhui WU, Xinxin LIU, Chen Y. Optimization and improvement of bond default early warning model - based on SMOTE-tomek-GWO-XGBoost method. *Friends Account* (2024)(06) 73–81. doi:10.3969/j.issn.1004-5937.2024.06.011

12. Fama EF, Jensen MC. Separation of ownership and control. *J L Econ* (1983) 26(2):301–25. doi:10.1086/467037

13. Wilson BA. Movements of wages over the business cycle: an intra-firm view. *Finance and Econ Discussion* (1997) 1997.0:1–38. doi:10.17016/feds.1997.01

14. Jun LU, Zifang LI. Empirical analysis of the effect of macro economic factors to corporate financial crisis. *J Shanxi Univ Finance Econ* (2008)(11) 94–100. doi:10.3969/j.issn.1007-9556.2008.11.017

15. Chen H. Macroeconomic conditions and the puzzles of credit spreads and capital structures. *The Natl Bur Econ Res* (2010) 65(6):2171–212. doi:10.1111/j.1540-6261.2010.01613.x

16. Xiao LIU, Rongxi ZHOU, Yuru LI. Default prediction of credit bond in China based on stacking algorithm integrated model. *Operations Res And Management Sci* (2023) 32(03):163–70. doi:10.12005/orms.2023.0096

17. Fan LZ, Zhang C. Explanation of Macro economic variables on bond risk premia in China. *J Management Sci China* (2009) 12(06):116–24. doi:10.3321/j.issn:1007-9807.2009.06.013

18. Dai guoqiang, Sun xinbao. On the macro determinants of credit spreads of corporate bonds in China. *J Finance Econ* (2011) 37(12):61–71.

19. Zhang K, Liang Z, Liu Z, Jia N. A deep learning based multitask model for network-wide traffic speed prediction. *Neurocomputing* (2020) 396:438–50. doi:10.1016/j.neucom.2018.10.097

20. Luo H, Cai J, Zhang K, Xie R, Liang Z. A multi-task deep learning model for short-term taxi demand forecasting considering spatiotemporal dependences. *J Traffic Transportation Eng (English Edition)* (2021) 8(1):83–94. doi:10.1016/j.jtte.2019.07.002

21. Yunjie W, Shaolong S, Kin Keung L, Ghulam A. A KELM-Based 587 ensemble learning approach for exchange rate forecasting. *J Syst Sci Infor* (2018) 6(4): 289–301. doi:10.21078/JSSI-2018-289-13

22. Xu F, Keselj V. Collective sentiment mining of microblogs in 24-hour stock price movement prediction. In: *16th IEEE conference on business informatics (CBI)* (2014). p. 60–7. doi:10.1109/CBI.2014.37

23. Almalaq B, Dai X, Zhang J, Hanrahan S, Nedrud J, Hebb A. Causality graph learning on cortical information flow in Parkinson's disease patients during behaviour tests. *49th Asilomar Conf Signals, Syst Comput* (2015) 925–9. doi:10.1109/ACSSC.2015.7421273

24. Yanwei H, Dengguo WU, Jun LI. Structural healthy monitoring data recovery based on extreme learning machine. *Computer Eng* (2011) 37(16):241–3. doi:10.3969/j.issn.1000-3428.2011.16.082

25. Vazquez-Patino A, Samaniego E, Campozano L, Avilés A. Effectiveness of causality-based predictor selection for statistical downscaling: a case study of rainfall in an Ecuadorian Andes basin. *Theor Appl Climatology* (2022) 150(3-4):987–1013. doi:10.1007/s00704-022-04205-2

26. Li M, Zhang R, Liu K. Machine learning incorporated with causal analysis for short-term prediction of sea ice. *Front Mar Sci* (2021) 8:8. doi:10.3389/fmars.2021.649378

27. Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases. *ACM SIGMOD Rec* (1993) 22(2):207–16. doi:10.1145/170036.170072

28. Schapire RE. The strength of weak learnability. *Machine Learn* (1990) 5(2):197–227. doi:10.1007/bf00116037

29. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. *Machine Learn* (2017). doi:10.48550/arXiv.1706.09516

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 31 March 2025
DOI 10.3389/fphy.2025.1550471

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Xiaoyun Xing,
Beijing Forestry University, China
Xueyong Liu,
Capital University of Economics and Business, China

*CORRESPONDENCE
Yiqiu Wang,
wangyiqiu@njau.edu.cn
Yang Xu,
xuyang@stu.njau.edu.cn

RECEIVED 23 December 2024
ACCEPTED 13 February 2025
PUBLISHED 31 March 2025

CITATION
Xu Y, Wang Y, Xie M and Guo C (2025) The impact of climate risk on the soundness of local commercial banks in China—based on physical risk and transition risk.
Front. Phys. 13:1550471.
doi: 10.3389/fphy.2025.1550471

COPYRIGHT
© 2025 Xu, Wang, Xie and Guo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The impact of climate risk on the soundness of local commercial banks in China—based on physical risk and transition risk

Yang Xu[1]*, Yiqiu Wang[1]*, Meng Xie[2] and Chong Guo[3]

[1]College of Finance, Nanjing Agricultural University, Nanjing, China, [2]School of Economics and Management, Henan Agricultural University, Zhengzhou, China, [3]College of Business, Suzhou Vocational University, Suzhou, China

Climate risk has a negative impact on the stable operation of social, economic, and financial systems. Local commercial banks operate in a single geographical area and are unable to effectively diversify risk. This article uses data from local commercial banks from 2009 to 2021 to explore the impact of climate risk on the soundness of local commercial banks. The research results show that climate risk, represented by extreme rainfall and transition risk, affects the soundness of local commercial banks by increasing direct losses to credit entities, reducing corporate profitability, and decreasing corporate credit demand. Further analysis shows that for rural commercial banks, extreme rainfall and extreme high temperatures diminish their soundness by affecting agricultural production. In regions with high temperatures, extreme heat reduces workers' labor income and the productivity of enterprises' labor, thereby affecting the soundness of local commercial banks. The larger the scale, the higher the degree of digital transformation, and the higher the level of property insurance of local commercial banks, the smaller the impact of climate risk on their soundness. Therefore, it is suggested that local commercial banks should raise awareness of climate risk, expand their business scope, increase insurance coverage for credit entities, and accelerate digital transformation to maintain stable operations in the context of increasing climate risk.

KEYWORDS

climate risk, local commercial banks, soundness, insurance, digitalization

1 Introduction

Climate change is a major challenge facing all mankind, seriously affecting the human living environment and socio-economic stability. As shown in Figure 1, the global warming trend is still ongoing, with frequent extreme high temperatures and an increase in extreme precipitation events in China. The frequency of extreme high-temperature events in China in 2022 was the highest since 1961 (Figure 2), and the climate risk index shows an upward trend. In order to reduce the impact of climate change risks in China, China proposed the "dual carbon target" set in September 2020, which aims to peak carbon dioxide emissions before 2030 and achieve carbon neutrality before 2060. In the 1990s, Schelling (1992), Ausubel (1991) [1, 2], and others began to discuss the impact of climate change on agriculture, economy, technological change, and other aspects.

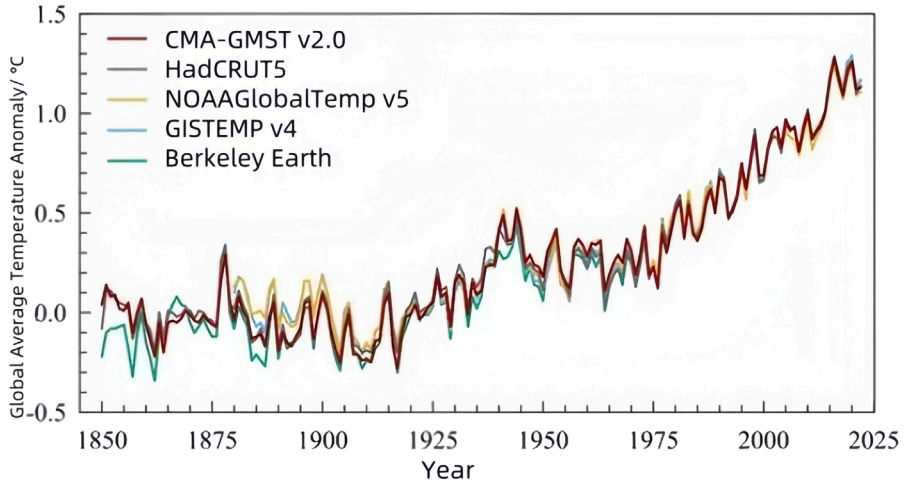


FIGURE 1
Global average temperature anomaly from 1850 to 2022 (relative to the average from 1850 to 1900). Reprinted with permission from Blue Book on Climate Change in China (2023) [90] by CMA Climate Change Centre, Copyright: Beijing, Science Press.

Carney (2015) [3] believes that climate change poses a threat to financial stability and long-term prosperity. With the implementation of the Paris Agreement, central banks and financial regulatory authorities of various countries have begun to study the possibility of financial systemic risks brought about by climate change. Climate change can trigger structural changes in the financial system, which has become a global consensus, but central banks around the world are divided over whether climate change factors should be included in the macro monetary policy and prudential supervision framework. The 20th National Congress of the Communist Party of China proposed actively participating in global governance to address climate change, strengthening the financial stability guarantee system, and safeguarding the bottom line of preventing systemic risks. The 2024 government work report also pointed out the vigorous development of the green and low-carbon economy, active and steady promotion of carbon peak neutrality, and classified disposal of financial risks. Climate change is increasingly being recognized as a contributing factor to financial risks.

The Bank of England's 2018 report found that most banks are starting to consider climate change risks like other risks, rather than simply a social responsibility. The People's Bank of China will carry out climate risk stress tests in 2021, organize 19 systemically important banks to conduct sensitivity stress tests on high-carbon industries, and improve financial institutions' risk management awareness. The European Central Bank stated that the losses incurred by European banks due to the rise in carbon asset prices and the impact of natural disasters exceeded €70 billion; The Bank of England found that banks and insurance companies that do not incorporate climate risk into their daily management may face profit losses of more than 10%. Climate risk has not yet been clearly recognized and effectively addressed by most financial institutions, and only a small number of large financial institutions adopt some environmental assessments (NGFS). Because small-scale banks are mostly local banks, their service recipients' information disclosure is not standardized and the quality of disclosure information is low. Therefore, small banks may face more severe climate risks (Liu et al., 2024) [4]. Local commercial banks include urban commercial banks, rural commercial banks, rural credit cooperatives, and village banks that operate within their residential areas. Compared to national banks, they cannot effectively diversify risks or avoid regional economic fluctuations (Wang et al., 2012) [5]. The business scope of national commercial banks and joint-stock commercial banks is more dispersed across regions and industries, and the impact of climate change on their operations is limited; Local commercial banks mainly serve counties with concentrated service areas. When extreme weather occurs, it has a significant impact on their operations (Liu et al., 2021; Pan et al., 2022) [6, 7]. Currently, China remains in an indirect financing system dominated by commercial banks. Local commercial banks undertake the mission of economic transformation, upgrading and development. Climate change increases the operational risks of banks, which are then transmitted throughout the financial system, triggering systemic risks. Therefore, discussing the impact of climate risks on the soundness of local commercial banks in the context of "dual carbon" is crucial for establishing a climate risk prevention framework for small- and medium-sized banks, as well as providing a basis for building a prudent regulatory system for climate risks, which holds practical significance.

The main marginal contributions of this article are as follows: (1) Existing literature focuses on the impact of climate risk on the credit risk of local commercial banks, but fails to reveal the negative effects of climate risk on the overall operation of local commercial banks. This article expands the single impact of climate risk on the credit risk of local commercial banks by examining the impact on their soundness. (2) This article makes a significant contribution to the literature by introducing a soundness indicator for commercial banks, which provides a more comprehensive measure of bank stability compared to the commonly used credit risk indicators in existing research. While prior studies have largely focused on the impact of climate risks on banks' risk-taking behavior using credit risk measures (Song et al., 2024) [8], this study extends the analysis by assessing the soundness of local commercial banks in China, incorporating both physical and transition risks. (3) By analyzing

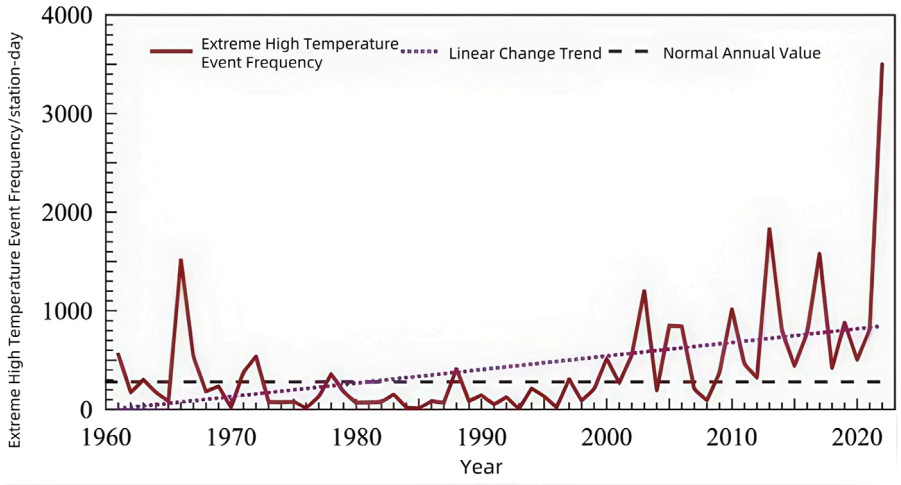


FIGURE 2
Frequency of Extreme High Temperature Events in China from 1961 to 2022. Reprinted with permission from Blue Book on Climate Change in China (2023) [90] by CMA Climate Change Centre, Copyright: Beijing, Science Press.

the heterogeneity of samples from multiple perspectives, this study explores the differences in the impact of different categories of physical risks on the heterogeneity of samples, and provides different climate risk compensation mechanisms for local commercial banks based on their own characteristics.

This study provides insights into the relationship between climate risks and bank soundness. However, we acknowledge that climate risks and financial regulations have continued to evolve beyond 2021. With the intensification of climate change, the frequency and intensity of extreme climate events are increasing, and financial institutions are continuously improving their risk management measures. Consequently, future research should incorporate more recent data to assess the dynamic nature of climate risks and the effectiveness of emerging regulatory policies (Campiglio et al., 2018) [9]. Additionally, new methodologies, such as stress testing and scenario analysis, can help model the impact of future climate risks on the financial sector more accurately (NGFS, 2020) [10]. The economic and financial aspects, as well as climate risk exposures, are closely interconnected among local commercial banks across different regions. While this study focuses on the impact of local climate risks on the soundness of individual local commercial banks, it does not fully account for the inter-regional synergistic effects. Future research could incorporate these cross-regional spillover effects to provide a more comprehensive understanding of how climate risks affect the banking sector across regions.

The remainder of this article is organized as follows: The second part reviews relevant literature on climate finance and banking operational risks, summarizes the transmission channels of climate risk to local commercial banks, and presents hypotheses. The third part outlines a research plan based on assumptions, presents the data, and constructs a regression model. The fourth part empirically examines the impact of regional climate risks on the soundness of local commercial banks and analyzes the results. The fifth part conducts a robustness test and further analysis to ensure the validity of the results. The sixth part summarizes the article and offers policy recommendations based on the experimental results.

2 Literature review and research hypotheses

Climate finance is an emerging research area in finance in recent years, such as the impact of climate risks on asset pricing, financial instruments, financial regulation, and public policy. The signing of the Paris Agreement in 2015 significantly advanced climate finance research. This chapter reviews the existing literature on climate finance and analyzes the mechanisms through which climate risk affects local commercial banks, proposing hypotheses.

The European Central Bank's 2019 Financial Stability Assessment Report noted that if climate risks are not reasonably priced by the market, they may pose systemic risks across the eurozone. Climate finance risks are characterized by the loss of underlying financial asset value due to climate change, with features of 'low probability and high loss'. Climate risk encompasses the risks from natural disasters, extreme weather, resource crises, and other negative impacts on human survival and economic growth due to climate change (Guo et al., 2023b) [11]. NGFS (2020) [10] also includes the risks associated with actions taken to address climate change within the scope of climate risk. Most researchers classify climate risk into physical and transition risks (Carney, 2015; NGFS, 2020; Cepni et al, 2023) [3, 10, 12]. Physical risks refer to direct economic losses from extreme weather events such as rainstorms, droughts, high temperatures, and floods, as well as indirect impacts caused by the ongoing effects of subsequent events; Transition risk refers to a series of policies introduced by various countries to address climate change, which lead to the accumulation of uncertain risks in the low-carbon transition market, such as the effects of policies, investor preferences, and market dynamics on economic and financial activities (Yang et al., 2024) [13]. Transition risk is long-term and typically occurs in industries with high levels of asset stranding, such as fossil energy (Zhang et al., 2022) [14].

Finance is the core and lifeblood of the modern economy. A financial crisis can have a profound impact on economic development. In China, banks occupy a central role in the financial

system. Initially, China established a highly centralized financial system dominated by large banks (Lin and Li, 2001) [15]. Yao and Dong (2014) [16] believe that China should establish a financial system centered around small and medium-sized banks. By the end of 2023, the total assets of China's banking industry are expected to account for over 90% of the total assets in financial institutions. This shows that China's financial system remains dominated by banks. Maintaining the stability of the banking sector is key to ensuring financial stability. Effective banking governance strengthens risk management, promotes stability, and is crucial to financial stability and economic growth (Wang, 2013) [17]. Systemic risks in banking can lead to economic crises and severe socio-economic consequences (Milne, 2009) [18]. A flawed banking system can reduce resource allocation efficiency and increase uncertainty in economic growth (Jokipii and Monnin, 2013) [19]. Stable operations are the fundamental principle of commercial banking, essential across different contexts and institutional conditions, and widely accepted internationally. Gong (2004) [20] suggests that sound banking operations should adapt to economic conditions and market changes, balance safety and efficiency, evolve steadily, and pursue development on a stable foundation. A bank's soundness reflects its ability to withstand market shocks. The lower the soundness, the more likely a bank is to face a run or collapse during systemic risk, which is linked to its fragility (Wang and Zhu, 2022) [21]. Chen and Wang (2011) [22] argue that existing research fails to fully capture the micro-level risks and robustness of bank operations, and propose a bank robustness indicator system based on five factors: capital adequacy, asset quality, profitability, liquidity, and risk. Most existing research examines the impact of climate change on commercial banks primarily from a risk perspective, often overlooking profitability and efficiency.

The primary source of commercial bank profits is the interest margin between deposits and loans, known as "interest income". Additionally, "non-interest income" comes from intermediary business activities (Wang and Liu, 2013) [23]. In the fourth quarter of 2023, non-interest income accounted for 19.93% of total income, while interest income exceeded 80%. This indicates that China's commercial banks still rely primarily on the traditional deposit-loan model for profits. Extreme weather events, such as high and low temperatures and heavy precipitation, can damage assets and reduce profitability for commercial bank borrowers, leading to direct or indirect losses for households and businesses, increasing the likelihood of defaults, and disrupting the stability of banks (Pan et al., 2022) [7]. Specifically, physical risks may lead to a variety of adverse financial outcomes for banks. For instance, extreme weather events such as floods, droughts, and storms can damage physical assets, disrupt supply chains, and reduce the productivity of key industries (e.g., agriculture, energy, manufacturing), all of which have significant financial implications for banks. When such risks are not adequately accounted for, they can lead to higher default rates on loans, especially in sectors most vulnerable to these events (e.g., agriculture, infrastructure, real estate) (Stern, 2007) [24]. Climate risks can cause corporate assets to become stranded or damaged, leading to fluctuations and depreciation in asset values, increasing uncertainty, and causing depreciation of assets mortgaged or held by commercial banks, thus increasing their operational risks (Ploeg and Rezai, 2020) [25]. Wu et al. (2024) [26] argue that the increase in climate risk will lead to a deterioration in the credit quality of commercial banks, resulting in heightened systemic risk, but with a limited impact on the depreciation of their investment portfolios. The intensification of climate risks has resulted in lower efficiency, business damage, supply chain disruptions, asset stranding, unstable income and profits, reduced repayment ability and willingness, rising default rates, and exacerbated bank credit risks, ultimately impacting their soundness. Climate risk not only reduces the profitability of enterprises but also intensifies financing constraints, lowers financing needs, reduces research and development investment, and ultimately decreases total factor productivity (Liu et al., 2023) [27]. Climate change significantly affects the operational efficiency of enterprises, leading to fixed asset losses, reduced expected returns, and decreased financial performance, including revenue, expenses, assets, liabilities, capital, and financing (Qing et al., 2024) [28]. In the context of banking regulation, the Basel III framework recognizes the importance of assessing and mitigating risk factors that could impair a bank's financial health. Several studies have emphasized the importance of integrating physical climate risks into stress testing and financial disclosure practices to improve the resilience of financial institutions (Stern, 2007) [24]. These insights suggest that integrating physical risks into the assessment of banking risk is essential to ensure the long-term stability and soundness of financial systems. In short, climate risks have increased financing constraints for businesses, reduced their financing capacity and demand, and led to a contraction in the credit scale of local banks. When bank deposit size remains unchanged, a decrease in loan size lowers the loan-to-deposit ratio. The loan-to-deposit ratio reflects a bank's asset allocation and is a key indicator of its operational efficiency. A higher loan-to-deposit ratio indicates greater efficiency in asset allocation (Zhen and Cao, 2005) [29]. Climate risk reduces corporate financing demand, decreasing the loan scale of local commercial banks, lowering their asset allocation efficiency, and further weakening their soundness.

In addition to the impact of physical risk, transition risk can also be transmitted through enterprises to banks. In order to alleviate the effects of climate change and extreme weather, the international community has begun advocating for low-carbon transformation and limiting carbon emissions, and has signed agreements such as the Paris Agreement to restrict member countries' carbon emissions. In the process of low-carbon transformation, financial risks have rapidly increased along with the accumulation of climate risks. Climate risks and natural disasters have led to instability in the banking industry, exacerbating financial risks (Hansen, 2022) [30]. In 2020, China proposed the "dual carbon" goal, striving to achieve carbon peak before 2030 and carbon neutrality before 2060. According to data from the International Energy Agency, China's carbon emissions in 2021 were 10.648 billion tons, accounting for 31.7%, ranking first in the world. China is facing enormous pressure to reduce emissions. Yu et al. (2022) [31] face severe challenges in emissions reduction, requiring significant changes in socio-economic development models and technologies, rapid development of low-carbon energy, decarbonization of energy transformation, and accelerated industrial restructuring. Thus, Chinese enterprises are facing enormous transformation challenges and regulatory pressures. Traditional environmental economic theory holds that environmental regulations affect the production conditions for optimal costs of enterprises. When the

internalization of environmental costs intensifies, it will squeeze out the funds invested by enterprises in other places, resulting in reduced efficiency, business damage, and decreased competitiveness in domestic and international markets (Yu and Sun, 2017) [32]. At the same time, strict environmental regulations have increased the financial risks of enterprises, which are more evident in terms of profitability and debt increase (Wang and Wang, 2024) [33]. In the process of transformation, high carbon emitting industries such as fossil steel face the risk of non clean energy assets being stranded due to emission restrictions (Sun et al., 2024) [34], which further leads to a decrease in the value of assets pledged to banks and increases their operational risks. Transition risks arise from significant changes in policy, technology, and market dynamics as economies shift towards sustainability. The transition to a low-carbon economy is driven by policies such as carbon pricing, regulations on emissions, and incentives for renewable energy. These policies directly impact financial markets, influencing capital flows and the cost of financing, particularly for carbon-intensive industries. According to Green Finance Theory, policies aimed at promoting green investments and sustainable practices redirect capital flows towards low-carbon technologies, which can influence the financial stability of banks with large exposures to high-carbon sectors (Lamperti et al., 2021) [35]. Furthermore, the financial transmission mechanisms explain how these policies affect banks through various channels. For instance, carbon taxes increase the cost of financing for carbon-intensive sectors, elevating credit risk for banks that are heavily invested in these industries. At the same time, the introduction of stricter environmental regulations can lead to asset devaluation in high-carbon industries, impacting the collateral value for loans (Carney, 2015; Battiston et al., 2017) [3, 36]. Moreover, the transition can also cause shifts in market expectations, creating both risks and opportunities for banks as they adjust their portfolios and risk management strategies (Nguyen et al., 2023) [37]. In addition, the interaction between policy uncertainty and financial markets is a critical factor in the financial transmission of climate transition risks. The introduction of new policies or changes in existing ones can cause uncertainty in the financial markets, leading to volatility in asset prices and shifting investment patterns (Pastor and Veronesi, 2012) [38]. Banks must navigate this uncertainty while adapting their risk assessment frameworks to mitigate the impact of potential policy changes on their financial soundness.

The above analysis indicates that both 'physical risks' and 'transition risks' affect the soundness of commercial banks through credit entities, potentially triggering systemic risks. Extreme weather events cause both direct and indirect economic losses to households and businesses. Over the past 50 years, extreme heat, low temperatures, precipitation, and other events have occurred frequently, reaching 11,778 cases, causing more than 2 million deaths and direct economic losses of up to 4.3 trillion US dollars. Given the increasing frequency of extreme weather events, direct and indirect losses caused by climate risks will be transmitted to the banking sector through credit entities, leading to higher default rates, lower collateral value, and profit losses, ultimately decreasing bank soundness. Based on this, the following hypotheses are proposed.

H1: Climate risks have negative impacts on the soundness of local commercial banks.

H2: Climate risk can cause direct economic losses to credit entities, reduce corporate profitability, increase corporate financing constraints, and reduce corporate financing demand, thereby reducing the soundness of local commercial banks.

Risk refers to the uncertainty surrounding future losses, including both the size and likelihood of such losses occurring (Guo et al., 2004) [39]. Huang et al. (2010) [40] defined meteorological disaster risk as a future scenario in which meteorological factors cause economic and social losses, based on the definitions of risk and natural disasters. The risk of natural disasters primarily affects the social economy in two ways: direct losses to the real economy and the anticipated impact of disasters on economic entities (Chen et al., 2014) [41]. Insurance has been highly successful in risk management, as it converts losses from uncertain or low-probability natural disasters into deterministic costs through premiums (Wang, 2006) [42]. Shao et al. (2017) [43] argue that the existence of insurance can reduce macroeconomic losses from risks and help mitigate economic and social fluctuations. For extreme weather, both pre- and post-disaster insurance measures can reduce losses to economic entities. Natural disasters are characterized by low probability and high loss, making the demand for natural disaster insurance by economic entities distinct from other forms of insurance. People tend to choose insurance for risks with a higher probability of occurrence. In China, natural disaster insurance has yet to play its full role, primarily due to insufficient demand for disaster insurance from economic entities, which remains far below global levels (Tian et al., 2015) [44]. Therefore, as climate risks intensify, insurance coverage becomes an effective means of mitigating losses for economic entities, thereby reducing the negative impact of credit entities on bank operations.

Digital technology disrupts enterprise production and operation by reducing costs, eliminating intermediaries, and transforming existing production models and organizational structures. Digital transformation has become a strategic goal for most banks. The digital transformation of commercial banks can significantly improve management, reduce operating costs, enhance operational efficiency, alleviate information asymmetry, and reduce risk-taking. The impact on small and medium-sized banks is more pronounced (Jiang et al., 2023; Zhai et al., 2023) [45, 46]. Improving the digitalization of local commercial banks strengthens risk management awareness, enhances understanding of climate risks, and reduces information asymmetry with high climate-risk credit entities. Digital transformation also helps commercial banks accurately identify enterprise credit needs, direct credit to physical businesses, prevent excessive corporate financialization, improve credit efficiency, and optimize credit structures (Ding and Zhou, 2024) [47]. Improving the digitalization of local commercial banks alleviates financing constraints, boosts financing needs,

expands credit volume, and enhances operational performance. Based on this, the following research hypotheses are proposed.

H3: The higher the level of insurance coverage and digitalization, the less impact climate risk has on the soundness of local commercial banks.

3 Data and methods

3.1 Data

This article uses data from local commercial banks in China from 2009 to 2021 as the research sample. Bank data are obtained from Wuhan University's China Banking Database (CBD), while climate data, including rainfall and temperature in China, is sourced from the publicly available FTP server of the National Climate Data Center (NCDC) in the United States. Enterprise and macroeconomic data are sourced from the WIND and CSMAR databases. The reason for choosing 2009 as the starting point is that, prior to 2009, the number of local commercial banks and the available data samples were relatively small and not representative. However, as of this study, data for 2022 and beyond have not yet been fully processed and standardized within the CBD. Therefore, the dataset currently extends only up to 2021, which remains highly representative of key trends in banking soundness and climate risk exposure. Although climate risks and financial policies continue to evolve, this period remains relevant for several reasons. First, it captures significant regulatory changes in climate-related financial risks, including the early implementation of green finance initiatives in China (Ma and Zadek, 2019) [48]. Second, it covers multiple extreme climate events, such as severe floods and heatwaves, allowing for an analysis of banks' responses to different climate risks (Song et al., 2024) [8]. Third, financial institutions have historically adapted their risk management strategies gradually, meaning that insights drawn from this period can still provide valuable guidance for understanding long-term trends in banking stability under climate risks (Bolton et al., 2020) [49]. To mitigate the influence of extreme values, a 1% truncation has been applied to bank financial variables. After data processing, the sample includes 293 local commercial banks, comprising 104 urban commercial banks, 185 rural commercial banks, and 4 rural credit cooperatives, with a total of 2,294 observations. The sample spans 31 provinces and autonomous regions, 131 prefecture-level cities, and 251 county-level administrative regions, demonstrating strong representativeness.

3.2 Research design

The study of climate risk typically requires controlling for time-invariant factors such as geographical characteristics. Therefore, using a fixed-effects model to analyze climate risk is more appropriate (Ding and Sun, 2022) [50]. To verify the above hypothesis, this article constructs the following baseline regression model for empirical analysis:

$$Sound_{i,t} = \beta_0 + \beta_1 Climate_{r,t} + \beta_2 Climate_{r,t-1} + \beta_3 X_{i,t}^{mic} + \beta_4 X_{r,t}^{mac} + \mu_i + \theta_t + \varepsilon_{i,t} \quad (1)$$

Among them, i represents a certain bank, t represents the year, and γ represents the prefecture level city where the bank is located. The dependent variable $Sound_{i,t}$ represents the operational robustness of bank i during period t. $Climate_{r,t}$ is the climate risk indicator for period t in region r.

Li and Li (2024) [51] believe that the risk impact of climate change on the insurance industry has a significant time lag. Fan et al. (2024) [52] found through a sample study of Chinese listed banks that climate risk has a delayed impact on banks. Similarly, there is a certain lag and persistence in the transmission of climate risks to banks, which means that previous climate risks may have an impact on current bank operations. Therefore, this article adds a first-order lag term of climate risk $Climate_{r,t-1}$ to the model. $X_{r,t}^{mic}$ represents the bank level control variable. $X_{i,t}^{mac}$ represents macro control variables at the regional level. μ_i is the individual fixed effect of the bank. θ_t is a time fixed effect. $\varepsilon_{i,t}$ is a random perturbation term.

3.3 Variable definition and explanation

3.3.1 Explained variable

Most scholars primarily consider bank stability—namely non-performing loans or risk-taking—when studying climate risks. However, the concept of bank soundness is distinct from other common bank risk indicators, as it focuses on a bank's ability to withstand various financial and environmental shocks. Unlike market, liquidity, or credit risk, which typically assess specific types of risks, soundness is a broader measure that integrates capital adequacy, liquidity management, and overall resilience to external pressures. Soundness is particularly valuable because it provides a comprehensive assessment of a bank's capacity to navigate financial turbulence or economic crises (Laeven and Levine, 2009; Coyle, 2000; Diamond and Dybvig, 1983) [53–55]. Soundness indicators offer a holistic evaluation, encompassing both short-term and long-term risks, making them ideal for assessing a bank's ability to endure diverse financial and market conditions. Additionally, soundness emphasizes a bank's preparedness for economic crises, ensuring that it is not only effectively managing risks but also structurally resilient over time (Basel Committee on Banking Supervision, 2011; Blundell-Wignall and Atkinson, 2010) [56, 57]. This makes soundness measures particularly relevant for evaluating both regulatory compliance and operational robustness across the banking sector. This article builds on Gong Feng's research (Gong, 2004) [20] and classifies bank soundness into three categories: stability indicators, development indicators, and efficiency indicators. Stability indicators include the capital adequacy ratio (Car), liquidity ratio (Liq), non-performing loan ratio (Npl), and provision coverage ratio (Pc). Development indicators comprise the asset growth rate (Dasset), profit growth rate (Dnrofit), and deposit growth rate (Ddeposit). Efficiency indicators consist of return on capital (Pcaptial), return on assets (Passet), income profit margin (Pincome), *per capita* profit (Pemp), and cost income rate (Cincome). This article employs the entropy method to synthesize 12 tertiary indicators into 3 secondary indicators, which are then consolidated into a single primary indicator. Since all indicators are standardized and normalized before applying the

entropy method, a higher bank soundness score indicates better overall bank performance.

3.3.2 Core explanatory variables

Climate risk is typically difficult to measure directly. This article draws on the research method of Pan et al. (2022) [7] and uses extreme weather events to quantify climate risk. Given the availability of extreme weather data, this article conducts empirical tests using extreme high temperatures (Htd), extreme low temperatures (Ltd.), and extreme rainfall (Ipd). Due to significant regional climate variations in China, this article employs the relative threshold method to determine extreme weather thresholds for each region. Temperature and rainfall data from 1981 to 2010 are used as the baseline for climate averages, and the extreme thresholds for local commercial banks are defined using daily meteorological observations from stations in the cities where banks are located. The percentile relative threshold method is applied to determine these thresholds, and the frequency of extreme events occurring each year during the sample period is calculated. Specifically, the methodology involves the following steps: Downloading data from the National Climate Data Center in the United States, extracting it, and selecting data from Chinese stations based on their country IDs. For temperature data, within each station, arranging the daily highest and lowest temperatures during the standard period from lowest to highest, and setting the 90th and 10th percentiles as the thresholds for extreme high and low temperatures, respectively. If, during the sample period, a station records a highest temperature exceeding the extreme high-temperature threshold or a lowest temperature falling below the extreme low-temperature threshold, an extreme high-temperature or low-temperature event is identified in the corresponding city. Daily data are then aggregated into annual reports to determine the number of extreme temperature days per year in each city. For rainfall data, sorting annual rainfall data (greater than 0 mm) within the standard period from lowest to highest, and setting the extreme rainfall threshold as the average of the 90th percentile of each year's rainfall data. Similarly, if a station records total annual rainfall exceeding this threshold during the sample period, an extreme rainfall event is identified in the city where the station is located. Daily data are then aggregated into annual reports to determine the number of extreme rainfall days per year in each city.

3.3.3 Control variables

To effectively identify the impact of climate risks, this article draws on the methodologies of Pan et al. (2022) [7], Liu et al. (2021) [6], and Xu (2022) [58] and selects the following control variables: Bank asset size (Lsize): The logarithm of total bank assets. Equity asset ratio (Ea): The proportion of equity assets to total assets. Deposit ratio (Dep): The proportion of total deposits to total assets. Net interest margin (Nim): The ratio of a bank's net interest income to its total interest-bearing assets. Macroeconomic variables are selected based on regional economic development and environmental conditions. Due to data availability, nighttime light brightness (Lights) is used as a proxy for economic development, while the average urban PM2.5 concentration is used as a proxy for environmental quality. The detailed definitions of the variables are presented in Table 1, and the descriptive statistics of the sample are shown in Table 2.

4 Empirical results

4.1 Analysis of baseline regression results

To test the impact of climate risk on bank soundness, this paper employs the regression model specified in Equation 1. The regression results are presented in Table 3. Columns (1)–(6) report the effects of extreme rainfall, extreme high temperatures, and extreme low temperatures on the soundness of local commercial banks. The results indicate a lagged effect of extreme rainfall on the soundness of local commercial banks. Current extreme rainfall does not significantly affect bank soundness, whereas one-period lagged extreme rainfall exhibit a significant negative impact. This negative effect remains statistically significant even after the inclusion of control variables. However, the effects of extreme high and extreme low temperatures remain insignificant, even after controlling for additional variables, suggesting that these temperature extremes do not have a meaningful impact on the soundness of local commercial banks. Ray et al. (2021) [59] found that the frequency of extreme events in India has increased over the past 50 years, with fatalities from floods far exceeding those caused by heatwaves and cold waves. Similarly, Hu et al. (2024) [60] assert that in China, extreme rainfall has the greatest impact and longest duration, persisting for up to 7 months. From an enterprise perspective, extreme rainfall increases environmental uncertainty and operational risks, exacerbating financing constraints, which hinders production activities and promotes financialization. However, extreme high and low temperatures do not exhibit significant effects (Deng et al., 2024) [61]. Extreme precipitation and extreme heat events are intensifying globally, with a rising likelihood of compound extreme events and low-probability, high-impact disasters (Zhai et al., 2021) [62]. In terms of disaster impact, economic losses, and post-disaster recovery, extreme rainfall tends to cause greater direct economic losses, affect a larger population, and require more complex and prolonged recovery efforts compared to extreme temperature events. According to an investigation by the Ministry of Emergency Management of the People's Republic of China, the consequences of extreme rainfall significantly outweigh those of extreme high temperatures. This suggests that compared to extreme high temperatures, extreme rainfall occurs more frequently, results in greater losses, affects a wider range, and presents greater challenges for disaster response and recovery. Therefore, the negative impact of extreme rainfall on the soundness of local commercial banks is significant, whereas the impact of extreme temperatures remains insignificant.

Climate risk can be categorized into physical risk and transition risk. The previous discussion focused on physical risks, represented by extreme temperatures and extreme rainfall. Transition risk refers to the uncertainty enterprises face in implementing green transformation under increasingly stringent climate policies. These risks include the stranding of high-carbon assets, energy substitution, technological innovation, and environmental regulations, all of which negatively affect business operations. Transition risk undoubtedly impacts the day-to-day operations of enterprises by raising production costs, such as expenses related to clean energy adoption and low-carbon innovation. These factors erode profitability and diminish asset values, which in turn transmit to the banking sector, ultimately

TABLE 1 Variable definition.

Category	Variable	Symbol	Definition
Explained variable	Soundness	Sound	Summarize 12 three-level indicators including stability, development, and efficiency using entropy method
Core explanatory variables	Climate Risk	Climate	Number of days of extreme events
	Extreme rainfall	Ipd	Number of days with extreme rainfall exceeding the threshold
	Extremely high- temperature	Htd	Number of days with extreme high-temperatures exceeding the threshold
	Extreme low- temperature	Ltd.	Number of days with extreme low-temperatures exceeding the threshold
	Transition risk	Translation	The proportion of GDP of high carbon emission industries to the GDP of prefecture level cities
Control variables	Bank asset size	Lsize	Logarithmic calculation of total assets
	Proportion of equity assets	Ea	Proportion of equity assets to total assets
	Deposit ratio	Dep	Deposit size as a percentage of total assets
	Net Interest Margin	Nim	The ratio of net interest income to total interest bearing assets of the bank
	Proportion of interest income	Irr	Net interest income to operating income ratio
	Night light brightness	Lights	Mean value of light points within the city where the bank is located
	PM2.5	PM2.5	Mean PM2.5 levels within the city where the bank is located

TABLE 2 Descriptive statistics of main variables.

Variable	Obs	Mean	Std	Min	Max
Sound	2,294	16.9589	5.3054	2.4895	43.7571
Ipd	2,294	13.9364	10.3183	0.0000	107.0000
Htd	2,294	60.7860	23.9786	0.0000	174.0000
Ltd.	2,294	31.5728	19.0100	0.0000	143.0000
Translation	670	59.7620	8.4540	25.5886	74.9238
Lsize	2,294	10.9234	1.3472	7.7617	13.8227
Ea	2,294	0.0795	0.0285	0.0061	0.9075
Dep	2,294	0.7446	0.1112	0.1819	0.9437
Nim	2,294	0.0324	0.0146	−0.0426	0.3957
Irr	2,294	0.7556	0.2351	−1.6747	2.6978
Lights	2,294	11.6298	9.7004	0.0415	56.6745
PM2.5	2,294	41.8800	15.3320	3.1912	138.0049

reducing the soundness of local commercial banks. Due to the lack of detailed data on commercial bank loan recipients, this article follows the methodology of Li and Pan (2022) [63] to measure transition risk at the prefecture-level city by using the proportion of GDP from high-carbon-emission industries relative to the city's total GDP. Specifically, we measure regional transition risk using the proportion of GDP from agriculture, forestry, animal husbandry and fishery, construction, real estate, industry, transportation, warehousing, and postal services. The empirical results are presented in Table 4. As shown in Table 4, transition risk also exerts a negative impact on the soundness of local commercial banks. This finding supports Hypothesis 1, confirming that climate risks have negative impacts on the soundness of local commercial banks.

4.2 Robust test

4.2.1 Adjusting thresholds and removing samples from areas with less impact of extreme events

To further verify the impact of extreme rainfall and extreme temperatures, we adjusted the thresholds for extreme rainfall, extreme high temperature, and extreme low temperature by both increasing and decreasing them. Specifically, the threshold for extreme rainfall was set at the 95th and 99th percentiles, the threshold for extreme high temperature at the 95th and 99th percentiles, and the threshold for extreme low temperature at the fifth and first percentiles. These stricter thresholds for extreme rainfall and temperature allow us to assess whether the regression results remain robust under different threshold specifications. The results are presented in Table 4.

Secondly, China has a vast geographic expanse and significant regional climate variations, leading to differences in both the

TABLE 3 Main regression results.

Variables	(1) Sound	(2) Sound	(3) Sound	(4) Sound	(5) Sound	(6) Sound
Ipd	−0.0021	0.0010				
	(-0.11)	(0.06)				
L.Ipd	−0.0400**	−0.0357*				
	(−2.14)	(−1.78)				
Htd			0.0091	0.0039		
			(0.69)	(0.81)		
L.Htd			0.0171	0.0111		
			(1.70)	(1.31)		
Ltd.					−0.0060	0.0024
					(−0.52)	(0.23)
L.Ltd.					−0.0052	0.0055
					(-0.45)	(0.55)
Observations	2,294	2,294	2,294	2,294	2,294	2,294
R-squared	0.098	0.136	0.100	0.136	0.096	0.135
Number of id	293	293	293	293	293	293
Control	NO	YES	NO	YES	NO	YES
Bank FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Note:***, **, * respectively indicate significance at the 1%, 5%, and 10% levels, with robust standard errors in parentheses, the same applies below.

TABLE 4 Impact of transition risk on the soundness of local commercial banks.

Variables	(1) Sound	(2) Sound
Translation	−0.1516**	−0.1155**
	(−2.18)	(−2.60)
Observations	670	670
R-squared	0.068	0.118
Number of id	273	273
Control	NO	YES
Bank FE	YES	YES
Year FE	YES	YES

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

probability and impact of extreme weather events. As this article employs the relative threshold method to measure extreme weather, the occurrence of extreme high temperatures in traditionally cooler regions and extreme low temperatures in warmer regions may not significantly affect social production and economic activities. This could introduce potential regression bias. Extreme temperature events in central and northern China exhibit strong long-range correlations, suggesting that extreme high temperature events are more likely to occur in these regions. Conversely, regions such as Yunnan-Guizhou, Northeast China, and the Qinghai-Tibet Plateau are less affected by extreme high temperatures (Ren et al., 2010) [64]. Similarly, extreme low temperatures have minimal impact on areas south of the Nanling Mountains. This paper follows the threshold criteria proposed by Cai (2016) [65] for extreme temperature events across different provinces. Specifically, provinces with an extreme high temperature threshold below 30°C and an extreme low temperature threshold above 0°C were excluded from the sample. This includes Guangdong, Guangxi, and Hainan, which

TABLE 5 Regression results after adjusting the threshold.

Variables Threshold	(1) Sound 95%	(2) Sound 99%	(3) Sound 95%	(4) Sound 99%	(5) Sound 5%	(6) Sound 1%
Ipd	0.0123	0.0171				
	(0.72)	(0.63)				
L.Ipd	−0.0368*	−0.0399*				
	(−1.77)	(−1.75)				
Htd			−0.0035	−0.0107		
			(−0.47)	(−0.99)		
L.Htd			0.0074	0.0008		
			(0.96)	(0.08)		
Ltd.					0.0011	−0.0003
					(0.09)	(−0.02)
L.Ltd.					0.0148	0.0352
					(1.06)	(1.53)
Observations	2,294	2,294	2,294	2,294	2,294	2,294
R-squared	0.136	0.135	0.135	0.135	0.135	0.136
Number of id	293	293	293	293	293	293
Control	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

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

are less affected by extreme low temperatures, as well as Qinghai and the Tibet Autonomous Region, which are less affected by extreme high temperatures. The results are presented in Table 5. In Table 5, Columns (1)–(2) present the results after excluding samples from Qinghai and Tibet, while Columns (3)–(4) show the results after excluding samples from Guangdong, Guangxi, and Hainan. The findings in Tables 5, 6 demonstrate that after adjusting the threshold criteria and removing specific samples, the results remain consistent with the baseline regression.

4.2.2 Replacing the standard period for extreme events and using other methods to measure climate risk

In the main regression, this article adopts the commonly used international practice of using the past 30 years as the standard period to calculate the extreme event threshold. Since the data in this article starts from 2009, the closest standard period is 1981–2010. In order to avoid the impact of incorrect selection of standard periods on empirical results, this article uses two standard periods, 1971–2000 and 1971–2010, to calculate extreme event thresholds and re conduct empirical tests. The regression results are shown in Table 7. According to the results in Table 7, after replacing the standard period, extreme rainfall still has a negative impact on the operational risk of local commercial banks, while extreme temperature does not have a significant impact, which is consistent with the main regression. The greater the fluctuation of temperature and rainfall, the greater the unexpected impact on the social economy. Therefore, this article draws on the approach of Liu et al. (2021) to measure climate risk using the fluctuation of annual average temperature and annual average rainfall, and conducts empirical testing again (We use temperature and rainfall fluctuations calculated based on the base period of 1991–2021). The specific results are shown in Table 8. The results in Table 8 show that rainfall fluctuations have a negative impact on the soundness of local commercial banks, and the impact lags behind; Temperature fluctuations have no significant impact on the soundness of local commercial banks, consistent with the main regression, demonstrating the robustness of the results.

4.2.3 Excluding the impact of special events

Since the reform and opening-up, China's rural credit cooperatives have undergone four stages of reform, with the period

TABLE 6 Regression results after excluding samples from areas with less impact of extreme events.

Variables	(1) Sound	(2) Sound	(3) Sound	(4) Sound
Ipd	0.0030		−0.0012	
	(0.17)		(-0.09)	
L.Ipd	−0.0341*		−0.0312*	
	(−1.73)		(−1.78)	
Htd		0.0030		
		(0.63)		
L.Htd		0.0105		
		(1.21)		
Ltd.				0.0051
				(0.44)
L.Ltd.				0.0036
				(0.35)
Observations	2,265	2,265	2,166	2,166
R-squared	0.143	0.142	0.137	0.135
Number of id	289	289	279	279
Control	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

from 2003 to 2010 marking the stage of provincial-level government management and comprehensive reform exploration (Sun and Wang, 2022) [66]. By the end of 2010, most of China's rural credit cooperatives had completed the transition to a shareholding system. To avoid the potential impact of these reforms on the analysis, we excluded samples prior to 2011. At the end of 2020, the outbreak of COVID-19 had a profound impact on China's economy and society. He et al. (2020) [67] argued that in the aftermath of COVID-19, the probability of enterprise defaults increased, significantly affecting urban and rural commercial banks. To eliminate potential biases caused by the COVID-19 shock, we also excluded samples after 2020. The results are presented in Table 9, which shows that the regression findings remain consistent with the main regression, demonstrating the robustness of the results.

4.2.4 Replace the explained variable

Soundness primarily emphasizes stability, which refers to a bank's ability and condition to maintain financial stability. Stability is the most critical factor in assessing bank soundness. According to Gong Feng's research (Gong, 2004) [20], the stability index score is the highest-weighted and most influential indicator of bank soundness. Therefore, we replaced the soundness indicator with the stability indicator for robustness testing and selected the default rate as a key variable for further analysis. The regression results, presented in Table 10, show that extreme rainfall has a negative impact on the stability of local commercial banks with a lag effect, while extreme high and low temperatures remain insignificant. Similarly, extreme rainfall significantly increases the non-performing loan ratio of local commercial banks without any lag, whereas extreme high and low temperatures have no significant impact on the non-performing loan ratio. Furthermore, transition risk significantly reduces the stability of local commercial banks, but does not have a significant impact on the non-performing loan ratio.

4.2.5 Endogeneity

To address the potential endogeneity issue between climate risk and the soundness of local commercial banks, we employed an instrumental variable (IV) approach. Specifically, we followed the methodology of Guo et al. (2023a) [68] and used the latitude of local commercial banks as an instrument for extreme rainfall. Latitude is an inherent geographical factor specific to each locality, and it satisfies the strict exogeneity condition required for instrumental variables. This choice is based on the fact that latitude is highly correlated with precipitation patterns, as regions closer to the equator typically receive more rainfall due to the higher levels of solar radiation and atmospheric convection that occur in these areas (Sauter et al., 2023) [69]. Latitude, as a geographical variable, is independent of the financial conditions of banks, making it an appropriate instrument for climate risk in the context of our study.

Table 11 presents findings from the instrumental variables' two-stage estimation. All the statistical figures derived from the Cragg-Donald Wald test surpass the critical threshold of 16.38, as determined at the 10% significance level by Stock-Yogo (2005), indicating that the problem of weak instrumental variables does not exist. The coefficients before Ipd in Columns (2) were negative and significant at the 10% significance level, which was consistent with the results of the main regression.

5 Further analysis

5.1 Mechanism analysis

The main credit targets of commercial banks are individuals, households, and enterprises. Climate risk affects the soundness of local commercial banks through three credit entities: individuals, households, and businesses. Extreme weather causes asset impairment and losses for bank credit entities, leading to a decrease in corporate profitability and increasing household income uncertainty. This affects individual, household, and corporate income, resulting in a decrease in the value of bank collateral assets, reduced repayment ability and willingness for individuals and businesses, a decline in the quality of commercial bank credit assets, an increase in default rates, a further increase in non-performing loans, and a decrease in profits. At the same time, from

10.3389/fphy.2025.1550471

TABLE 7 Regression results after replacing the standard period for extreme events.

Variables Period	(1) Sound	(2) Sound	(3) Sound	(4) Sound	(5) Sound	(6) Sound
	1971–2000			1971–2010		
Ipd	0.0027			0.0023		
	(0.15)			(0.12)		
L.Ipd	−0.0388*			−0.0368*		
	(−1.79)			(−1.75)		
Htd		0.0049			0.0061	
		(0.96)			(1.24)	
L.Htd		0.0129			0.0121	
		(1.67)			(1.47)	
Ltd.			0.0027			0.0017
			(0.23)			(0.15)
L.Ltd.			0.0081			0.0086
			(0.86)			(0.86)
Observations	2,264	2,264	2,264	2,294	2,294	2,294
R-squared	0.135	0.135	0.133	0.137	0.136	0.135
Number of id	290	290	290	293	293	293
Control	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

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

the perspective of profit and risk, banks need to raise interest rates to compensate for the increased operational risks of enterprises, which increases the financing costs of enterprises, increases financing constraints, reduces financing demand, and leads to a reduction in bank credit supply. Therefore, the impact of climate risk on the soundness of local commercial banks is mainly through the direct losses caused by extreme weather to credit subjects, the indirect effects on the profitability of enterprises, and the reduction of credit supply.

5.1.1 Direct economic losses incurred by credit entities

This article uses the logarithm of the direct economic losses caused by natural disasters at the provincial level published by the National Bureau of Statistics as an intermediary variable to measure the direct economic losses of commercial bank service providers. Referring to the mediation effect model proposed by Wen and Ye (2014) [70], a stepwise test method was used to construct a model to examine whether climate risk affects the soundness of local commercial banks by affecting the direct economic losses of service providers in Equation 2 and in Equation 3.

$$Loss_{r,t} = \alpha_0 + \alpha_1 Climate_{r,t} + \alpha_2 Controls_{r,t} + \mu_r + \theta_t + \varepsilon_{r,t} \quad (2)$$

$$Sound_{i,t} = \beta_0 + \beta_1 Loss_{r,t} + \beta_2 Loss_{r,t-1} + \beta_3 Climate_{r,t} + \beta_4 Climate_{r,t-1} + \beta_5 X_{i,t}^{mic} + \beta_6 X_{r,t}^{mac} + \mu_i + \theta_t + \varepsilon_{i,t} \quad (3)$$

The regression results are shown in Supplementary Table S1. Columns (2) and (3) in Supplementary Table S1 report the regression results of models (2) and (3). The Ipd coefficient in column (2) is positive and significant at the 1% significance level, indicating that extreme rainfall can cause direct economic losses to bank service recipients. In column (3), L The loss coefficient is negative and significant, with Ipd and L Ipd is not significant, indicating that the direct economic losses of bank service recipients have a negative impact on the soundness of local commercial banks, and this impact has a lag effect. The results of columns (1)–(3) indicate that the direct economic losses of service providers are the

TABLE 8 Regression results after replacing explanatory variables.

Variables	(1) Sound	(2) Sound	(3) Sound	(4) Sound
Prec	−0.1737	−0.2413		
	(−0.79)	(−1.29)		
L.Prec	−0.7992***	−0.7010***		
	(−3.85)	(−2.95)		
Temp			−0.1050	−0.3323
			(−0.42)	(−1.38)
L.Temp			0.0036	−0.1172
			(0.02)	(−0.54)
Observations	1,184	1,133	2,307	2,224
R-squared	0.147	0.200	0.096	0.140
Number of id	259	253	281	275
Control	NO	YES	NO	YES
Bank FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

TABLE 9 Regression results excluding special events.

Variables	(1) Sound	(2) Sound	(3) Sound	(4) Sound
Ipd	−0.0038			
	(−0.29)			
L.Ipd	−0.0593*			
	(−1.93)			
Htd		0.0018		
		(0.32)		
L.Htd		0.0077		
		(0.82)		
Ltd.			−0.0043	
			(−0.37)	
L.Ltd.			0.0021	
			(0.18)	
Translation				−0.0888**
				(−2.14)
Observations	1,893	1,893	1,893	526
R-squared	0.162	0.159	0.158	0.087
Number of id	291	291	291	270
Control	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

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

mediating variable for the impact of climate risk on the soundness of local commercial banks.

5.1.2 Decreased profitability of enterprises

Climate risk worsens the profitability of enterprises, affects their production status, increases the probability of default, and thus affects the soundness of local commercial banks. Due to the inability to obtain bank level enterprise data, it is not possible to conduct bank enterprise matching in a simple way. Therefore, this article refers to the approach of Bertrand and Mullainathan (2001) [71] and Ge et al. (2021) [72], selects a sample of A-share listed companies, and establishes the following two-stage model to explore the impact mechanism of climate risk on the soundness of local commercial banks - whether the profitability of enterprises is established:

$$\boldsymbol{FrimRisk}_{i,t} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{Climate}_{r,t} + \boldsymbol{\beta}_2 \boldsymbol{X}_{i,t}^{mic} + \boldsymbol{\beta}_3 \boldsymbol{X}_{r,t}^{mac} + \boldsymbol{\mu}_i + \boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_{i,t} \quad (4)$$

$FrimRisk_{i,t}$ represents the profitability of a company, and $X_{i,t}^{mic}$ and $X_{r,t}^{mac}$ represent control variables at the company and regional levels, respectively. The coefficient before β_1 represents the impact of climate risk on the profitability of listed companies. This article uses operating profit/total assets to measure the profitability of enterprises, and controls for variables such as the logarithm of total assets, survival turnover rate, industry competition level (Herfindahl index), regional GDP growth rate, Tobin's Q value, and Hu Huanyong line as dummy variables for regional economic characteristics. Due to the fixed individual and time effects in Equation 4, the estimated value of β_1 represents the average impact of climate risk after excluding individual and macro dimensional influences. This article sets up a time region dimension indicator to connect the Equation 4:

$$\boldsymbol{\hat{FrimRisk}}_{r,t} = \hat{\boldsymbol{\beta}}_1 \boldsymbol{Climate}_{r,t} \quad (5)$$

$$\boldsymbol{Sound}_{i,t} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{\hat{FrimRisk}}_{r,t} + \boldsymbol{\beta}_2 \boldsymbol{\hat{FrimRisk}}_{r,t-1} + \boldsymbol{\beta}_3 \boldsymbol{X}_{i,t}^{mic} + \boldsymbol{\beta}_4 \boldsymbol{X}_{r,t}^{mac} + \boldsymbol{\mu}_i + \boldsymbol{\theta}_t + \boldsymbol{\varepsilon}_{i,t} \quad (6)$$

The estimated coefficients in Equation 5 multiplied by $Climate_{r,t}$ represent the average impact of climate risk on businesses in the region. Replace the estimated value of $\hat{FrimRisk}_{r,t}$ with the values of $Climate_{r,t}$ and $Climate_{r,t-1}$ in Equation 1 to obtain the second stage Equation 6. The coefficient β_1 represents how the changes in corporate operational capacity caused by climate risks in the region

TABLE 10 Regression results for replacing the dependent variable.

Variables	(1) Stability	(2) Stability	(3) Stability	(4) Stability	(5) Npl	(6) Npl	(7) Npl	(8) Npl
Ipd	−0.0157				−0.0025**			
	(−0.45)				(−2.50)			
L.Ipd	−0.0719*				−0.0003			
	(−1.84)				(−0.32)			
Htd		−0.0045				−0.0005		
		(−0.38)				(−0.98)		
L.Htd		0.0098				−0.0003		
		(0.58)				(−0.52)		
Ltd.			−0.0000				−0.0012	
			(−0.00)				(−1.60)	
L.Ltd.			0.0025				−0.0011	
			(0.12)				(−1.39)	
Translation				−0.2150**				0.0000
				(−2.29)				(0.01)
Observations	2,294	2,294	2,294	670	2,986	2,986	2,986	828
R-squared	0.160	0.158	0.158	0.126	0.200	0.198	0.202	0.173
Number of id	293	293	293	273	354	354	354	341
Control	YES	YES	YES	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

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

TABLE 11 Instrumental variable (IV) estimation.

Variables	(1) First stage L.Ipd	(2) Second stage Sound
IV	−0.4680***	
	(−16.79)	
L.Ipd		−0.1039**
		(−2.36)
Observations	2,315	2,315
Control	YES	YES
Cragg-Donald Wald		281.803

Note: Column (2) and column (4) are Z statistics in brackets.
*, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

affect the soundness of bank. The results of the two-stage regression are shown in Supplementary Table S2. According to the results of the first stage in Supplementary Table S2, climate risks represented by extreme rainfall have reduced the profitability of enterprises. Based on the comprehensive regression results of the two stages, it can be concluded that climate risks can have a negative impact on the soundness of local commercial banks by reducing their profitability.

5.1.3 Corporate financing constraints, bank credit scale

Climate risk increases the financing constraints for enterprises, reduces their financing needs, leads to a decrease in operational efficiency, increases financial difficulties, and reduces the credit scale of local commercial banks, thereby affecting soundness. This article selects the logarithm of the total loan amount of local commercial banks and the enterprise SA index as intermediary variables to measure the scale of bank loans and financing constraints of

enterprises, referring to the mediation effect model of Wen and Ye (2014) [70] and the two-stage model in the previous section.

$Loan_{i,t}$ represents the logarithm of the total loan amount, and $SA_{i,t}$ represents the financing constraint of the enterprise. The regression results are shown in Supplementary Table S3. The results in columns (1)–(2) of Supplementary Table S3 show that extreme rainfall can significantly reduce the credit scale of local commercial banks, leading to a decrease in the soundness of banks. The regression results of the two-stage model are shown in columns (3)–(4) of Supplementary Table S3. The results indicate that extreme rainfall exacerbates the financing constraints of A-share listed companies. By constructing an average impact index of climate risk on financing constraints of enterprises in various regions, it can be concluded that extreme rainfall exacerbates financing constraints, leading to a decrease in financing demand and an increase in financing difficulty, which in turn affects the soundness of local commercial banks.

5.2 Heterogeneity analysis

The heterogeneity analysis is crucial in understanding the varying impacts of climate risks on local commercial banks. Different types of banks (such as state-owned or private banks) may respond differently to these risks due to varying levels of government support and risk management capacities (Yeung, 2021) [73]. Moreover, larger banks, with their more extensive resources, are better equipped to handle climate risks compared to smaller institutions (Battiston et al., 2017) [36]. Finally, geographical differences play a key role as regions with higher exposure to physical climate risks, such as coastal areas or arid regions, experience more severe impacts, which can affect their financial stability (Jia et al., 2022) [74].

In order to capture the impact of transition risks on local commercial banks, we acknowledge that enterprises' technological innovation capabilities and policy adaptation capabilities play a crucial role. However, due to data limitations, particularly the absence of detailed reports on technological innovation and policy adaptation capabilities for many banks in our sample, we were unable to directly incorporate these factors into our analysis. As an alternative, we examined bank size and bank nature (rural commercial banks vs urban commercial banks), as they are often linked to the technological innovation and policy adaptation capacities of the enterprises within the banks' portfolios. Research has shown that larger banks tend to have more diversified portfolios and better access to technological innovation, which enables them to better manage transition risks (Zhu, 2013) [75]. Similarly, rural commercial banks tend to be more affected by policies, and their policy adaptability is weak, thus weakening their capacity to navigate transition risks (Hao, 2022) [76]. By analyzing these characteristics, we indirectly assess the potential influence of technological innovation and policy adaptation capabilities on the climate risk exposure of the banks.

5.2.1 Nature of commercial bank

Agricultural production is one of the most sensitive sectors to climate change, and climate is the primary factor affecting agriculture. Climate change has led to significant changes in China's crop planting system and production layout, and the increase in climate risks will directly affect the agricultural economy, resulting in a decrease in crop yields (Ding and Sun, 2022) [50]. Meanwhile, there is an inverted U-shaped relationship between precipitation and crop production. Extreme precipitation leads to waterlogging in farmland, loss of soil fertility, and affects crop yield, which is a factor contributing to the vulnerability of agricultural production (Zhou et al., 2010) [77]. Liu et al. (2012) [78] studied the impact of extreme weather on agricultural economic output in China. The research results showed that climate risk is the Granger cause of reduced agricultural economic output, and extreme rainfall has the greatest impact on agricultural economy. Due to the different nature of local commercial banks, there are significant differences in their service providers and business scope. Rural commercial banks mainly serve agricultural production oriented objects such as agriculture, rural areas, and farmers, and their business scope is mostly related to agriculture, rural areas, and farmers, with geographical locations concentrated in rural areas such as counties. We divided the sample into rural commercial banks (RCB) and urban commercial banks (UCB) for grouped regression to test whether the nature of banks affects the operational robustness of local commercial banks under climate risk. The regression results are shown in Supplementary Table S4, which indicates that neither extreme rainfall nor extreme temperature have a significant impact on the soundness of urban commercial banks. Extreme rainfall has no significant impact on the soundness of rural commercial banks at the 90% threshold, but when the threshold is raised, the impact of extreme rainfall on their operations significantly increases, indicating that the higher the degree of extreme rainfall, the greater the impact on the soundness of rural commercial banks. The impact of extreme high temperatures on the soundness of rural commercial banks increases with the increase of the high temperature threshold, and there is no lag effect. In the basic regression, extreme high temperatures have no impact on the operation of local commercial banks, but due to the significant negative impact of high temperatures on agricultural production, extreme high temperatures have a significant negative impact on rural commercial banks serving agriculture, rural areas, and farmers. There is no significant impact of extreme low temperature on the soundness of rural commercial banks under the three thresholds. The possible reason is that the average annual temperature in China has been increasing year by year, with a significant warming in winter and a relatively small impact of low temperatures on crop yields (Pan et al., 2011) [79]. Yu (2019) [80] studied the impact of climate change on crop production in Zhejiang Province. The study showed that extreme low temperatures had limited effects on crop yield reduction, with the highest impact on rapeseed yield reduction. For every day of increase in extreme low temperatures, rapeseed yield decreased by only 0.53%–0.79%, while extreme high temperatures had a greater impact on crop yield. For every day of increase in extreme high temperatures, early rice yield decreased by 3.91%–5.08%. Therefore, under the background of climate change, the impact of extreme low temperatures on agricultural production and economic and social development is limited, and rural commercial banks are not significantly affected by it. This article also conducts heterogeneity analysis based on the nature of commercial banks on transition risk. The results are shown in Supplementary Table S5.

According to Supplementary Table S5, transition risk also have a greater impact on the soundness of rural commercial banks. Rural commercial banks tend to have weaker resistance to policy risks compared to larger banks. This is primarily due to their limited capacity to adapt to policy changes and lower human resource investments. These banks rely heavily on government support and are more vulnerable to policy shifts, especially those related to agricultural financing. Additionally, their lack of advanced risk management and technological resources limits their ability to handle transition risks effectively. Consequently, rural commercial banks face greater challenges and more significant impacts on their financial stability when dealing with transition risks, as they have fewer resources and less capacity to manage such risks compared to larger, more well-equipped banks.

5.2.2 Scale of commercial banks

The larger the scale of a bank, the wider its business scope, the more service objects it serves, and the more branches it has. Its assets are relatively diversified, and risks are more easily spread across various industries. Jiang and Chen (2012) [81] believe that the scale of commercial banks has the function of risk diversification. The larger the scale, the more diversified the operating assets, which can effectively diversify certain risks and reduce their risk-taking. This is more conducive to maintaining soundness and preventing systemic risks from occurring. This article starts from the perspective of local commercial bank assets, and divides the sample of local commercial banks into large and small banks through asset grouping regression to determine the impact of climate risk on the soundness of local commercial banks of different sizes. The results are shown in Supplementary Table S6. According to Supplementary Table S6, extreme rainfall has a negative impact on the soundness of smaller local commercial banks, and the impact increases with the increase of the threshold, while the impact on larger local commercial banks is not significant. The impact of extreme high and low temperatures on the soundness of new commercial banks in both large and small areas is not significant at different thresholds. The results in Supplementary Table S7 indicate that transition risk also has a negative impact on the soundness of smaller local commercial banks. Larger banks generally possess a significant advantage in technological innovation and human capital compared to smaller institutions like rural commercial banks. Their ability to invest in advanced technologies and specialized talent enhances their capacity to manage transition risks effectively. This enables them to adapt more efficiently to policy shifts and market changes. As a result, larger banks experience a lower impact from transition risks, with their financial stability and resilience being less affected compared to smaller banks. Both physical and transformational risks have a greater impact on smaller local commercial banks, making it difficult to effectively diversify risks.

5.2.3 Commercial bank region

With the intensification of global climate change, high temperatures have caused a series of impacts on human health. Temperature can increase negative emotions in humans, leading to psychological disorders. Yu et al. (2020) [82] believe that meteorology has an emotional effect, and teenagers, the elderly, and workers are susceptible to high temperatures and high temperature stress. KO et al. (2024) [83] analyzed data from 140 countries and found that climate change significantly reduces the happiness of people in the region. The impact of extreme high temperatures on labor productivity shows an inverted U-shaped relationship. When the temperature is too high, it will reduce the labor productivity of workers. Due to the decrease in labor productivity of workers, enterprises will increase mechanized production to replace labor (Huang et al., 2023; Wang and Lin, 2021) [84, 85]. Therefore, based on Cai's research [65], this article retains samples from provinces with high temperature thresholds above 36 °C to empirically test the impact of extreme high temperatures on the soundness of local commercial banks. The regression results are shown in Supplementary Table S8. Extreme high temperatures only reduce the soundness of local commercial banks at the 99% threshold, indicating that in high-temperature areas, the operational robustness of local commercial banks will only be affected when extreme temperatures further intensify.

5.3 Moderating effect

5.3.1 The mitigating effect of insurance on climate risks

Based on the above analysis, climate risk can cause direct economic losses to individuals, households, and businesses, resulting in fixed asset losses, reduced income, decreased repayment ability and willingness, and a decline in the soundness of local commercial banks. Insurance can convert uncertain losses and risks into fixed premiums, and convert uncertain losses faced by individuals, households, and businesses into fixed costs, greatly enhancing their ability to resist risks. Therefore, purchasing property insurance can enhance one's ability to resist risks and prevent low-risk, high loss risk shocks at a lower cost. This article uses property insurance income data from various provinces to divide the samples into high and low levels of property insurance, and conducts empirical tests by grouping them. Meanwhile, using the interaction term approach, the moderating effect of property insurance was verified, and the results are shown in Supplementary Table S8. According to Supplementary Table S8, climate risk has a negative impact on the soundness of local commercial banks in samples with low insurance levels, while the impact is not significant in samples with high insurance levels. At the same time, under the interaction term model, the coefficient before the interaction term between extreme rainfall and property insurance income is positive, indicating that property insurance has a positive moderating effect on the impact of climate risk on the operational robustness of local commercial banks, which is consistent with hypothesis 3.

Climate has a significant impact on agricultural production, and rural commercial banks are greatly affected by climate risks. But insurance can mitigate the negative impact of climate risks on the operation of local commercial banks. Agricultural insurance can supplement farmers' risk diversification ability in large-scale natural disasters and help restore production (Zhang et al., 2007) [86]. Agricultural insurance can alleviate the impact of climate risks on agricultural production and the risks transmitted to local commercial banks. This article uses agricultural insurance income data from various provinces to empirically test whether the level of agricultural insurance protection can alleviate the impact of climate risks on the soundness of local commercial

banks. As shown in Supplementary Table S9, agricultural insurance is beneficial in mitigating the negative impact of climate risks on local commercial banks. But the coefficient is not significant under the interaction term model, which may be due to the relatively backward and narrow coverage of China's agricultural insurance system in the early days. In 2013, China implemented the "Regulations on Agricultural Insurance", and the development of agricultural insurance entered a new stage (We excluded samples before 2014, and the coefficient before the interaction term was significantly positive, proving the impact of agricultural insurance on mitigating climate risks).

Regarding the transition risk, this article also divided the samples into two groups based on property insurance and agricultural insurance, with high and low insurance levels, for grouped regression. The regression results are shown in Supplementary Table S10. The results in Supplementary Table S10 show that the levels of agricultural insurance and property insurance have no mitigating effect on transition risks. The possible reason is that transition risk is the risk of increased costs and decreased profits in the production process due to environmental regulations, carbon emission restrictions, and other factors. Therefore, purchasing property insurance and agricultural insurance cannot have a positive effect on environmental penalties, asset obsolescence, and green technology innovation faced by enterprises during the transformation process.

5.3.2 The mitigating effect of digitalization on climate risks

The digital transformation of local commercial banks has the effect of reducing costs and increasing efficiency, effectively alleviating the degree of information asymmetry between banks and enterprises, improving operational efficiency, and expanding credit scale. According to the above analysis, improving the digitalization level of local commercial banks is beneficial for mitigating the negative impact of climate risks on their soundness. This article uses an interaction term model to verify the effect of digital level mitigation of climate risk on the operational robustness of local commercial banks. The measurement of digitalization level is generally obtained by using Python text analysis to aggregate the frequency of keywords appearing in the "digitalization" dictionary in the annual reports of listed companies (Wu et al., 2021) [87]. Local commercial banks are different from listed companies in that it is difficult to collect annual reports and disclose information irregularly, making it difficult to obtain effective digital transformation keywords. This article draws on the approach of Qi et al. (2024) [88] and uses a "Python text analysis + information retrieval" method to construct a digitalization level index for local commercial banks, measuring the level of digitalization. Python text analysis are used to statistically analyze keyword search results, which are summarized by bank year. This index has obvious right bias, and we perform logarithmic processing on it. The regression results are shown in Supplementary Table S11. The significant improvement in digitalization has alleviated the impact of extreme rainfall on the soundness of local commercial banks. Meanwhile, the risk of transformation does not have a significant impact on the soundness of local commercial banks with high levels of digital transformation, but still has a significant negative impact on the soundness of local commercial banks with low levels of digital transformation.

6 Conclusion and suggestions

This article uses a sample of 293 local commercial banks from 2009 to 2021 to categorize climate risk into physical risk (extreme high temperatures, extreme low temperatures, and extreme rainfall) and transition risk. It quantitatively analyzes the impact of climate risk on the soundness of local commercial banks and tests the heterogeneity of this impact across bank size, type, and property insurance levels. The main conclusions of the empirical test are as follows: (1) Climate risk, represented by extreme rainfall, significantly reduces the soundness of local commercial banks, while extreme high and low temperatures have no significant impact across the entire sample. (2) Climate risk affects the soundness of local commercial banks by impacting the direct losses and profitability of individuals, households, and businesses, reducing repayment capacity and loan willingness. (3) Agriculture is an industry highly impacted by climate risks. Extreme high temperatures and extreme rainfall negatively affect the soundness of rural commercial banks, with the impact intensifying as climate risks increase, but no significant effect on urban commercial banks. (4) Large-scale local commercial banks have stronger risk diversification capabilities, while extreme rainfall has a greater impact on smaller-scale local commercial banks, with the effect increasing as climate risk intensifies. (5) Property insurance preserves assets at low cost and converts uncertain climate risks into premiums. (6) Digital transformation reduces information asymmetry among local commercial banks, improves capital utilization, and mitigates the impact of climate risks on their soundness. (7) Both transition and physical risks negatively impact the soundness of local commercial banks. Due to differences in impact channels, transition risk has a greater effect on rural and smaller banks, while the alleviating effect of agricultural insurance is not significant.

Based on the above analysis, this article proposes the following recommendations:

First, raise awareness of climate risks among local commercial banks and integrate them into their daily management practices. Different types of climate risks should be assigned appropriate weights. For risks such as extreme rainfall, which can lead to significant losses, preventive measures and contingency plans should be established in advance.

Second, rural commercial banks with a high proportion of agricultural credit should carefully consider the impact of local climate on agricultural production, including whether specific agricultural products are affected by extreme temperatures and rainfall. In daily operations, when collateral is involved, it is crucial to assess the potential for climate risks to cause damage, and for collateral vulnerable to climate risks, its value should be appropriately adjusted.

Third, local commercial banks should pursue diversified business operations to broaden their scope and mitigate climate risks. A concentrated business scope makes operations highly vulnerable to climate risks, which can have a significant impact. Diversifying into other industries less affected by climate risks

can help mitigate these risks and enhance operational stability. Additionally, risk hedging products can be developed to manage potential climate risks. Certain agricultural products are less impacted by extreme temperatures and rainfall, with some crops even benefiting from these conditions. Credit supply can be adjusted based on the type of operation to hedge against climate risks.

Fourth, local commercial banks should encourage or even mandate property insurance for individuals facing higher climate risks, incorporating insurance purchase decisions into credit assessments to determine credit limits and interest rates. When granting credit to individual and family customers, banks should consider their insurance coverage to ensure sufficient risk protection. Banks can collaborate with insurance companies to offer policy incentives for high-quality customers, improving bank soundness and boosting insurance companies' revenues. Simultaneously, local commercial banks should accelerate their digital transformation, aiming to reach higher levels of digitalization, reduce information asymmetry between banks and enterprises, and enhance fund utilization efficiency.

Finally, the feasibility of digital transformation varies across banks due to differences in financial resources, technological investment capabilities, and talent availability (Cuesta et al., 2015) [89]. Large commercial banks have the resources to invest in advanced AI-driven risk prediction models, while small and medium-sized banks may face challenges due to limited technical expertise and funding constraints. To effectively integrate digital transformation into climate risk management, banks can adopt a phased implementation strategy based on their size and technological readiness: Large banks: Can establish dedicated climate risk management departments, develop AI-driven climate risk models, and integrate blockchain for secure green finance transactions. Medium-sized banks: Can focus on partnerships with fintech firms, adopting cloud-based climate risk analysis tools, and enhancing data-sharing platforms for climate finance. Small banks: Should prioritize cost-effective solutions, such as utilizing government-provided climate risk databases, outsourcing data analysis to specialized firms, and leveraging public-private collaborations for technology adoption.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

YX: Data curation, Methodology, Software, Writing–original draft, Writing–review and editing. YW: Conceptualization, Funding acquisition, Methodology, Writing–original draft. MX: Data curation, Writing–review and editing. CG: Investigation, Visualization, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. Major Project of Philosophy and Social Science Research of Guizhou Province (No. 22GZZB07) and Humanities and Social Science Fundamental Research Funds for Central Universities of the Nanjing Agricultural University (No. SKCX2023013).

Acknowledgments

We sincerely thank Professor Yan Liu from Wuhan University for providing the data. All errors are authors' own.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1550471/full#supplementary-material

References

1. Schelling TC. Some economics of global warming. *Am Econ Rev* (1992) 82:1–14.

2. Ausubel JH. Does climate still matter? *Nature (London)* (1991) 350:649–52. doi:10.1038/350649a0

3. Carney M. *Breaking the tragedy of the horizon – climate change and financial stability*. London: Speech at Lloyd's of London (2015). Available from: https://www.bankofengland.co.uk/speech/2015/breaking-the-tragedy-of-the-horizon-climate- change-and-financial-stability.

4. Liu Z, Li J, Sun H. Climate transition risk and bank risk-taking: the role of digital transformation. *Financ Res Lett* (2024) 61:105028. doi:10.1016/j.frl.2024.105028

5. Wang Q, Wu W, Huang J. Geographic diversification of city commercial banks: lending expansion, risk level and bank performance. *J Financial Res* (2012) 01:141–53.

6. Liu B, Wang XH, Li MX. Climate change and the credit risk of rural financial institutions. *J Financial Res* (2021) 12:96–115.

7. Pan M, Liu HY, Cheng ZS. The impact of extreme climate on commercial banks' risk-taking: evidence from local commercial banks in China. *J Financial Res* (2022) 10:39–57.

8. Song Y, Chen W, Zhang K. The impact of climate change on the risk-taking level of Chinese commercial banks: empirical evidence from Chinese local commercial banks. *Clim Change Econ* (2024) 15:2440002. doi:10.1142/s2010007824400025

9. Campiglio E, Dafermos Y, Monnin P, Ryan-Collins J, Schotten G, Tanaka M. Climate change challenges for central banks and financial regulators. *Nat Clim Chang* (2018) 8:462–8. doi:10.1038/s41558-018-0175-0

10. NGFS (Network for Greening the Financial System). Guide to climate scenario analysis for central banks and supervisors (2020). Available from: https://www.ngfs.net.

11. Guo K, Li Y, Zhang Y, Ji Q, Zhao W. How are climate risk shocks connected to agricultural markets? *J Commod Mark* (2023) 32:100367. doi:10.1016/j.jcomm.2023.100367

12. Cepni O, Demirer R, Pham L, Rognone L. Climate uncertainty and information transmissions across the conventional and esg assets. *J Int Financial Markets, Institutions and Money* (2023) 83:101730. doi:10.1016/j.intfin.2022.101730

13. Yang ZH, Li DC, Chen YT. Research on the "green swan" risk in financial markets: based on the dual perspectives of physical risk and transition risk. *J Manage World* (2024) 40(02):47–67. doi:10.19744/j.cnki.11-1235/f.2024.0016

14. Zhang S, Lu LP, Zhang XM, Wang H. Climate risk in financial system: assessment, pricing, and management. *Chin Rev Financial Stud* (2022) 14(01):99–120.

15. Lin YF, Li YJ. Promoting the growth of medium and small-sized enterprises through the development of medium and small-sized financial institutions. *Econ Res J* (2001) 1:11–8.

16. Yao YJ, Dong GF. The development of small and medium-sized banks and the financing constraints of SMEs: empirical research from the perspective of optima financial structure theory in new structural economics. *J Finance Econ* (2014) 40(01):105–15. doi:10.16538/j.cnki.jfe.2014.01.010

17. Wang M. *A research on the impact of shocks between China's banking soundness and macro-economic*. Jilin: Jilin University (2013).

18. Milne A. "The fall of the house of credit: borrowing short and lending long: the illusion of liquidity in structured credit," in *The Fall of the House of Credit*. Cambridge: Cambridge University Press. (2009). doi:10.1017/CBO9780511596643.007

19. Jokipii T, Monnin P. The impact of banking sector stability on the real economy. *J Int Money Finance* (2013) 32:1–16. doi:10.1016/j.jimonfin.2012.02.008

20. Gong F. Research on the soundness operation of China's banking industry under the background of financial globalization. Shanghai: *Shanghai Acad Soc Sci* (2004).

21. Wang H, Zhu JY. Bank soundness and liquid asset allocation from the perspective of financial regulation. *Econ Res J* (2022) 57(12):104–23.

22. Chen SD, Wang M. Study on stability of China's banking system. *J Quantitative and Technol Econ* (2011) 28(10):64–77. doi:10.13653/j.cnki.jqte.2011.10.008

23. Wang MS, Liu XF. Research on the impact of income structure of commercial banks on profitability: analysis based on panel data of 14 listed banks in China. *Nankai Business Rev* (2013) 16(02):143–9. doi:10.3969/j.issn.1008-3448.2013.02.015

24. Stern N. *The economics of climate change: the stern review*. Cambridge University Press (2007).

25. Ploeg RVD, Rezai A. Stranded assets in the transition to a carbon-free economy. *Annual Review of Resource Economics* (2020) 12 (1). doi:10.1146/annurev-resource-110519-040938

26. Wu B, Wen F, Zhang Y, Huang ZJ. Climate risk and the systemic risk of banks: a global perspective. *J Int Financial Markets, Institutions and Money* (2024) 95:102030. doi:10.1016/j.intfin.2024.102030

27. Liu B, Li YQ, Jiang YJ. Climate changes and the total factor productivity of manufacturing enterprises: mechanism analysis and empirical tests. *J Hunan Univ (Social Sciences)* (2023) 37(01):78–87. doi:10.16339/j.cnki.hdxbskb.2023.01.011

28. Qing L, Li P, Dagestani AA, Woo C, Zhong K. Does climate change exposure impact on corporate finance and energy performance? Unraveling the moderating role of ceos' green experience. *J Clean Prod* (2024) 461:142653. doi:10.1016/j.jclepro.2024.142653

29. Zheng LJ, Cao TQ. Empirical analysis of efficiency and influencing factors of commercial banks in China. *J Financial Res* (2005) 1:91–101.

30. Hansen LP. Central banking challenges posed by uncertain climate change and natural disasters. *J Monet Econ* (2022) 125:1–15. doi:10.1016/j.jmoneco.2021.09.010

31. Yu GR, Hao TX, Zhu JX. Discussion on action strategies of China's carbon peak and carbon neutrality. *Bull Chin Acad Sci* (2022) 37(04):423–34. doi:10.16418/j.issn.1000-3045.20220121001

32. Yu DH, Sun T. Environmental regulation, skill premium and international competitiveness of manufacturing industry. *China Ind Econ* (2017) 05:35–53. doi:10.19581/j.cnki.ciejournal.2017.05.004

33. Wang J, Wang H. Environmental regulation and corporate financial risk: the role of credit guarantees. *Financ Res Lett* (2024) 63:105357. doi:10.1016/j.frl.2024.105357

34. Sun J, Yang Y, Zhou P. Low-carbon transition risks in the energy sector: a systematic review. *Sustain Prod Consum* (2024) 50:115–27. doi:10.1016/j.spc.2024.07.025

35. Lamperti F, Bosetti V, Roventini A, Tavoni M, Treibich T. Three green financial policies to address climate risks. *J Financial Stab* (2021) 54:100875. doi:10.1016/j.jfs.2021.100875

36. Battiston S, Mandel A, Monasterolo I, Schütze F, Visentin G. A climate stress-test of the financial system. *Nat Clim Chang* (2017) 7:283–8. doi:10.1038/nclimate3255

37. Nguyen Q, Diaz-Rainey I, Kuruppuarachchi D, McCarten M, Tan EK. Climate transition risk in US loan portfolios: are all banks the same? *Int Rev Financial Anal* (2023) 85:102401. doi:10.1016/j.irfa.2022.102401

38. Pastor L, Veronesi P. Uncertainty about government policy and stock prices. *J Finance* (2012) 67(4):1219–64. doi:10.1111/j.1540-6261.2012.01746.x

39. Guo XT, Pu YJ, Lin L. Risk concept and its quantitative characterization. *J Quantitative and Technol Econ* (2004) 21(2):111–5. doi:10.3969/j.issn.1000-3894.2004.02.018

40. Huang CF, Liu AN, Wang Y. A discussion on basic definition of disaster risk. *J Nat Disasters* (2010) 19(6):8–16. doi:10.13577/j.jnd.2010.0602

41. Chen GJ, Zhao JF, Wu XL, Zhao XQ. Rare disaster risks and macroeconomic fluctuations in China. *Econ Res J* (2014) 49(08):54–66.

42. Wang GJ (2006). Insurance economics. Insurance economics.

43. Shao QQ, Wang B, Bai LF. Risk shocks, insurance elements and economic fluctuations in China. *J Financial Res* (2017) 06:1–16.

44. Tian L, Yao P, Wang HB. The relationship among government behavior, risk perception and catastrophe insurance demand. *China Soft Sci* (2015) 09:70–81. doi:10.3969/j.issn.1002-9753.2015.09.007

45. Jiang H, Tang SF, Wu WY. Research on the lmpact of digital transformation on commercial banks risk-taking-theoretical logic and empirical evidence. *Stud Int Finance* (2023) 01:62–73. doi:10.16475/j.cnki.1006-1029.2023.01.006

46. Zhai SB, Cheng YT, Xie L. Impact of commercial banks' digital transformation on their risk-taking level. *J Beijing Technol Business Univ (Social Sciences)* (2023) 38(02):75–86. doi:10.12085/j.issn.1009-6116.2023.02.007

47. Ding X, Zhou Y. Digital transformation and bank credit allocation: based on the perspective of bank lending to the real economy. *J Quantitative and Technol Econ* (2024) 41(03):193–216. doi:10.13653/j.cnki.jqte.20240004.001

48. Ma J, Zadek S. *Decarbonizing the belt and road: a green finance roadmap*. Beijing: Tsinghua Univ. and UNEP Rep (2019).

49. Bolton P, Despres M, Pereira da Silva LA, Samama F, Svartzman R. The green swan. *BIS Books, Bank for International Settlements, number 31* (2020).

50. Ding YG, Sun QX. Impact of climate risks on China's agricultural economic development: heterogeneity and mechanism analyses. *J Financial Res* (2022) 09:111–31.

51. Li ZH, Li KH. The theoretical analysis and empirical testing of the risk impact of climate change on the insurance industry. *INSURANCE STUDIES* (2024) 07:18–35. doi:10.13497/j.cnki.is.2024.07.002

52. Fan WN, Wang F, Zhang H, Ling R, Jiang HF. Heterogeneous responses of commercial banks to various climate risks: evidence from 42 a-share listed banks in China. *Clim Chang Econ (Singap)* (2024) 15. doi:10.1142/S2010007824500064

53. Laeven L, Levine R. Bank governance, regulation, and risk-taking. *J Financ Econ* (2009) 93:259–75. doi:10.1016/j.jfineco.2008.09.003

54. Coyle B. *A practical guide to risk management*. New Jersey: Prentice Hall (2000).

55. Diamond DW, Dybvig PH. Bank runs, deposit insurance, and liquidity. *J Polit Econ* (1983) 91:401–19. doi:10.1086/261155

56. Basel Committee on Banking Supervision. Basel III: a global regulatory framework for more resilient banks and banking systems. *Bank Int Settl Rep* (2011). Available from: https://www.bis.org/publ/bcbs189.htm (Accessed October 15, 2024).

57. Blundell-Wignall A, Atkinson P. The current financial crisis: causes and policy issues. *OECD J Financ Mark Trends* (2010) 1:87–121. doi:10.1787/fmt-v2010-1-en

58. Xu HX. *Research on the impact of climate change on bank credit risk*. Wuhan: Zhongnan University of Economics and Law (2022). doi:10.27660/d.cnki.gzczu.2022.001770

59. Ray K, Giri RK, Ray SS, Dimri AP, Rajeevan M. An assessment of long-term changes in mortalities due to extreme weather events in India: a study of 50 years' data, 1970–2019. *Weather Clim Extrem* (2021) 32:100315. doi:10.1016/j.wace.2021.100315

60. Hu L, Meng J, Xiong C, Fang W, Yang J, Liu M, et al. City-level resilience to extreme weather shocks revealed by satellite nighttime lights in China. *Sustain Cities Soc* (2024) 101:105167. doi:10.1016/j.scs.2023.105167

61. Deng C, Su Z, Feng Y. Extreme climate and corporate financialization: evidence from China. *Econ Anal Pol* (2024) 81:306–21. doi:10.1016/j.eap.2023.12.001

62. Zhai PM, Zhou BQ, Chen Y, Yu R. Several new understandings in the climate change science. *Clim Change Res* (2021) 17(6):629–35. doi:10.12006/j.issn.1673-1719.2021.201

63. Li S, Pan Z. Climate transition risk and bank performance: evidence from China. *J Environ Manage* (2022) 323:116275. doi:10.1016/j.jenvman.2022.116275

64. Ren GY, Feng GL, Yan ZW. Progresses in observation studies of climate extremes and changes in mainland China. *Climatic Environ Res* (2010) 15(04):337–53.

65. Cai WX. *Temporal trends features and spatial distributions of extreme climate events in China*. Beijing: University of International Business and Economics (2016).

66. Sun XF, Wang CC. Research on the impact of shareholding reform of rural credit cooperatives on county economic growth in China. *J Finance Econ* (2022) 48(4):154–68. doi:10.16538/j.cnki.jfe.20211218.402

67. He CY, Wen YC, Chang YL, Gen XX. Measurement and analysis of the COVID-19 epidemic impact on China's economy. *J Quantitative and Technol Econ* (2020) 37(5):3–22. doi:10.13653/j.cnki.jqte.2020.05.001

68. Guo K, Bian Y, Zhang DY, Ji Q. The impact of climate-related risk on local government financing costs. *J Environ Econ* (2023) 8(03):108–31. doi:10.19511/j.cnki.jee.2023.03.006

69. Sauter C, Fowler H, Westra S, Ali A, Peleg N, White C. Compound extreme hourly rainfall preconditioned by heatwaves most likely in the mid-latitudes. *Weather Clim Extremes* (2023) 40:100563. doi:10.1016/j.wace.2023.100563

70. Wen ZL, Ye BJ. Analyses of mediating effects: the development of methods and models. *Adv Psychol Sci* (2014) 22(05):731–45. doi:10.3724/SP.J.1042.2014.00731

71. Bertrand M, Mullainathan S. Are CEOs rewarded for luck? The ones without principals are. *Q J Econ* (2001) 116(3):901–32. doi:10.1162/00335530152466269

72. Ge XY, Zhuang JL, Liu Y. How does trade policy uncertainty affect bank risk—an examination of the firm performance channel. *China Ind Econ* (2021) 8:133–51. doi:10.3969/j.issn.1006-480X.2021.08.008

73. Yeung G. Chinese state-owned commercial banks in reform: inefficient and yet credible and functional? *J Chin Governance* (2021) 6(2):198–231. doi:10.1080/23812346.2020.1772537

74. Jia S, Yang C, Wang M, Failler P. Heterogeneous impact of land-use on climate change: study from a spatial perspective. *Front Environ Sci* (2022) 10:840603. doi:10.3389/fenvs.2022.840603

75. Zhu MX. *Research on the formation mechanism of China's listed commercial banks' innovative capability and its influence on bank performance*. Shandong: Shandong University (2013).

76. Hao Z. Risk sources of small and medium sized commercial banks in China. *Doctoral dissertation, ISCTE-Instituto Universitario de Lisboa Portugal* (2022).

77. Zhou SD, Zhou WK, Zhu GH, Wang CX, Wang Y. Impact of climate change on agriculture and its countermeasures. *J Nanjing Agric Univ (Social Sci Edition)* (2010) 10(01):34–9. doi:10.3969/j.issn.1671-7465.2010.01.006

78. Liu J, Xu XF, Luo H. Empirical study on the impact of extreme weather and climate events on agricultural economic output in China. *Scientia Sinica (Terrae)* (2012) 42(7):1076–82.

79. Pan GX, Gao M, Hu GH, et al. Issues and challenges on mitigation of climate change impacts on China's future agriculture. *J Agro-Environment Sci* (2011) 30(9):1707–12.

80. Yu SA. *The impact of climate change on crop production. An empirical study in Zhejiang*. Zhejiang: Zhejiang University (2019). doi:10.27461/d.cnki.gzjdx.2019.000671

81. Jiang SX, Chen YC. Monetary policy, bank capital and risk-taking. *J Financial Res* (2012)(04) 1–16.

82. Yu GL, Chen TT, Zhao FQ. The influence of air temperature and temperature variability on menta health. *Adv Psychol Sci* (2020) 28(8):1282–92. doi:10.3724/SP.J.1042.2020.01282

83. Ko J, Leung CK, Chen X, Palmer DA. From emissions to emotions: exploring the impact of climate change on happiness across 140 countries. *Glob Transitions* (2024) 6:231–40. doi:10.1016/j.glt.2024.10.005

84. Huang QH. *The impact of extreme temperature to labor income: evidence from Chinese manufacturing firms Emission-Water saving-Employment Impact in China*. Shandong: Shandong University (2023).

85. Wang CC, Lin QQ. How does bad weather affect labor productivity? — Evidence from the adaptation behavior of couriers. *China Econ Q* (2021) 21(03):797–818. doi:10.13821/j.cnki.ceq.2021.03.03

86. Zhang YH, Shi QH, Gu HY. A theoretical and positive study on the demand of crop insurance. *J Quantitative and Technol Econ* (2007) 24(4):65–75. doi:10.3969/j.issn.1000-3894.2007.04.007

87. Wu F, Hu HZ, Lin HY, Ren XY. Enterprise digital transformation and capital market performance: empirical evidence from stock liquidity. *J Manage World* (2021) 37(07):130–44. doi:10.19744/j.cnki.11-1235/f.2021.0097

88. Qi H, Zhou YS, He LF. Digital transformation and the improvement of the dual performance of rural financial institutions. *China Rural Surv* (2024) 1:108–26. doi:10.20074/j.cnki.11-3586/f.2024.01.006

89. Cuesta C, Ruesta M, Tuesta D, Urbiola P. The digital transformation of the banking industry. *BBVA Res* (2015) 1(4):1–10. doi:10.1109/TMAG.2013.2243424

90. CMA Climate Change Centre. Blue Book on Climate Change in China (2023). Beijing: Science Press (2023).

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 15 April 2025
DOI 10.3389/fphy.2025.1578995

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Quanpeng Chen,
Sichuan Vocational College of Finance and Economics, China
Shaojian Qu,
Nanjing University of Information Science and Technology, China

*CORRESPONDENCE
Qiaoming Hou,
houqiaoming@sut.edu.cn

RECEIVED 18 February 2025
ACCEPTED 06 March 2025
PUBLISHED 15 April 2025

CITATION
Yang L and Hou Q (2025) Research on bank financing strategy of e-commerce supply chain driven by blockchain.
Front. Phys. 13:1578995.
doi: 10.3389/fphy.2025.1578995

COPYRIGHT
© 2025 Yang and Hou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Research on bank financing strategy of e-commerce supply chain driven by blockchain

Lei Yang and Qiaoming Hou*

School of Management, Shenyang University of Technology, Shenyang, China

Introduction: In the e-commerce supply chain, many risk-averse suppliers face the dilemma of capital constraints. Bank financing helps to alleviate these financing difficulties, and blockchain trust mechanisms can optimize bank financing strategies.

Methods: To address the issue of capital constraints for suppliers, based on the two operating modes of Consignment and Direct Selling in the e-commerce supply chain, using the Stackelberg game, considering the company's risk-averse behavior and consumers' green preferences, a game model is constructed to explore the impact of factors such as blockchain, risk aversion, product greenness, guarantee interest rates, and information verification efforts on bank financing strategies.

Results and discussion: When the blockchain technology is adopted within a certain threshold, the supplier can obtain higher financing income by adopting consideration than Direct Selling. Compared with the Direct Selling mode, the blockchain trust mechanism under consideration mode can play a better role and improve the financing efficiency of suppliers. The financing efficiency of suppliers will also decrease with the increase of risk aversion coefficient, and the impact of risk aversion on suppliers will be greater in the case of blockchain, but the trust of blockchain will increase with the increase of risk aversion coefficient. Product greenness will increase the sales price of e-commerce supply chain under consideration mode, and improving product greenness will help improve enterprise cash flow. However, under the Direct Selling mode, the impact on the price is not great, and the impact on the financing strategy will be different. Although the guaranteed interest rate will increase the financing cost, it will make banks more willing to provide financing for suppliers with financial constraints. In practice, managers should comprehensively consider the balance between guaranteed interest rate and financing return. Information verification efforts will increase product greenness and green preference coefficient, and the improvement of information transparency will make e-commerce platforms more willing to provide guarantees for suppliers. This study is expected to provide decision-making references for the managers of e-commerce supply chain.

KEYWORDS

blockchain, risk aversion, green preference, e-commerce supply chain, bank financing

1 Introduction

The e-commerce market is growing rapidly, and the market size of e-commerce supply chains is expanding, leading to an increasing demand for Supply Chain Financing (SCF) and growing concerns about financing constraints. According to the "2023 Global E-commerce Platform Report" published by Webretailer, global green e-commerce sales $6.3 trillion, with an annual growth rate of 9% in 2023. It is expected to reach $8.1 trillion by 2025. The continued growth of the green e-commerce market has made the operation of e-commerce green supply chains a key area of focus [1]. Consumer green preferences are driving suppliers to focus more on producing green and low-carbon products. To meet market demands, e-commerce supply chains have adopted two sales models: Direct Selling and Consignment. In the Direct Selling model, which is platform-led, inventory funding is high, and financing demand is rapidly increasing. In the Consignment model, which is supplier-led, the operation cycle is longer, and cash flow is insufficient. With market expansion, the demand for funding in e-commerce supply chains continues to grow. According to Market Research Future, the global supply chain financing demand was valued at $235 billion in 2023. According to the World Bank's report on "Financing Gaps for Small and Medium Enterprises," the financing demand of Small and Medium-sized Enterprises (SMEs) accounted for 40% of global supply chain financing demand in 2023. The financing needs of e-commerce supply chains are enormous, but existing funds cannot fully meet operational requirements, with about 35% of businesses facing severe financing constraints.

Bank financing is the most widely used financing model, and finance strategies can address the financing constraints of e-commerce green supply chains. In bank financing, suppliers with financial constraints are guaranteed by the e-commerce platform and apply for loans from banks to solve their own financial problems [2]. For funding-constrained suppliers, bank finance is preferred due to its simple process and comprehensive audit measures [3]. In the bank financing model, higher interest rates increase financing costs, and seller competition is detrimental to improving financing efficiency. Higher product pricing can generate more cash flow, better alleviating funding constraints [4]. Operating capital levels also affect financing strategy formulation. A higher guaranteed interest rate encourages e-commerce platforms to provide more guarantees. The platform's service fee rate should not be too high, as this could affect the financing willingness of funding-constrained firms [5]. Both interest rates and product pricing affect financing costs, and the platform's sales model also influences financing strategies, making it crucial to choose the right sales model to improve financing efficiency.

In e-commerce supply chain bank financing, there are many risk-related issues, and businesses are increasingly focusing on risk aversion. Bank financing relies on e-commerce platforms to provide guarantees for suppliers to resolve their funding constraints. However, suppliers are often SMEs with low credit ratings and significant credit and moral risks [6]. Information asymmetry between enterprises leads to barriers in the finance review process, and the information obtained cannot guarantee its authenticity, resulting in a lack of trust mechanisms and the risk of fraudulent transaction data [7]. E-commerce supply chain enterprises are generally risk-averse, and risk-averse suppliers tend to adopt conservative, preventive measures in financing decisions to reduce potential risks and ensure operational stability [8]. Risk-averse e-commerce platforms assess supplier credit by reviewing the transaction data within the e-commerce supply chain, a data-based risk control approach that allows the platform to effectively reduce potential risks and meet its own risk aversion preferences [9].

Blockchain technology holds great potential for improving bank financing strategies in e-commerce supply chains. Blockchain's trust mechanism enhances the credibility between enterprises, reduces the risks faced during financing, and improves the financing efficiency of suppliers. According to Grand View Research, the application market for blockchain in e-commerce green supply chain financing is expected to reach $26.4 billion by 2027, primarily benefiting from the transparency blockchain provides, which reduces fraud risk, alleviates financial constraints, and improves liquidity. Blockchain automates the recording of transaction data, reducing the cost of information collection and enabling real-time dynamic credit ratings based on blockchain data. Especially in multi-level e-commerce supply chains, blockchain's trust mechanism better facilitates trust transmission, reducing credit risks and moral hazards related to adverse selection. The SAS platform of Ping An Bank uses the distributed ledger technology of blockchain to transfer and manage accounts receivable, which solves the problem of information sharing and trust transfer between transaction parties, and improves the financing efficiency of banks. In various power structures within the e-commerce supply chain, blockchain can better protect the interests of followers. According to JD.com's statistics, JD Finance utilizes blockchain to enhance supply chain financial services, shortening the financing cycle by 30%–40% and reducing financing costs by 15%–25%. Blockchain enhances trust in financing data, reduces reliance on intermediaries, simplifies financing processes, lowers financing costs, and increases financing returns [10]. The "Double Chain Link" platform launched by Ant Financial takes advantage of the immutability and transparency of blockchain to enhance financial institutions' trust in transaction information, reduce financing costs and risk levels, and optimize bank financing strategies.

To solve the funding constraints in e-commerce supply chains and better help enterprises avoid risks, this study uses blockchain technology to optimize bank financing strategies and improve financing efficiency. The study constructs a bank financing model for e-commerce supply chains before and after adopting blockchain, considering Direct Selling and Consignment sales models, and explores how blockchain and risk aversion affect bank financing strategies. The study addresses the following key questions: (1) Will bank financing differ under different power structures in the e-commerce supply chains with Direct Selling and Consignment models? How can the sales model be chosen to seek the optimal bank financing strategy? (2) What is the difference between the bank financing strategies of e-commerce supply chain before and after the adoption of blockchain? How to optimize bank financing strategy by using blockchain trust mechanism? The contributions are as follows: First, from the perspective of blockchain trust, it discusses the impact of trust between enterprises and between enterprises and the market on bank financing strategies, providing new insights into the study of factors affecting bank financing. Second, using game theory analysis, it investigates the impact of enterprises' risk aversion on bank financing strategies before and after blockchain adoption, finding that higher risk aversion does not necessarily lead to improved financing efficiency, which contradicts common

assumptions. Finally, Consignment the Direct Selling and consideration of e-commerce supply chain, this paper explores the choice of sales methods to seek the best bank financing strategy. The traditional two-level e-commerce supply chain is extended to a three-level deep-level e-commerce supply chain, and the bank financing strategies under different power structures of the supply chain are analyzed. With different sales methods and power structures, the bank financing strategies will be different, and appropriate sales methods should be chosen to improve financing efficiency.

The structure is as follows: Section 2 summarizes existing literature. Section 3 is the practical issues of the research, defines relevant symbols, and presents research hypotheses. Section 4 builds three models: the baseline bank financing model, risk-averse e-commerce green supply chain bank finance model, and blockchain-enabled risk-averse e-commerce green supply chain bank finance model, followed by comparative analysis. Section 5 conducts numerical analysis of the models using MATLAB, examining the effects of blockchain, blockchain trust, risk aversion, product greenness, guarantee interest rates, and information collection efforts on e-commerce supply chain bank financing strategies. Section 6 concludes the paper, discussing its limitations and proposing future directions.

2 Literature review

2.1 Bank financing strategies in E-commerce supply chain

Research on banking financing strategies in e-commerce supply chains mainly focuses on financing influencing factors, financing risks, financing strategy models, and strategy optimization.

Regarding the influencing factors of financing, interest financing impacts the willingness of constrained supply chains to take loans. Supply chain companies can achieve Pareto optimality and better financing benefits when financing rates fall within a certain threshold. When the rate is smaller, the order quantity increases, and the expected profit for online retailers also rises. Additionally, initial working capital also affects the choice of financing strategies [11]. Under the Direct Selling model, when the unit costs of products with different qualities are the same, and the differences in e-commerce service fees and product evaluations are minimal, manufacturers constrained by funds are more likely to opt for bank financing [12]. Other scholars have analyzed various factors that influence bank financing, exploring the operational mechanisms and internal influences of SCF. They found that government subsidies and third-party guarantees also affect e-commerce supply chain bank financing strategies [13]. The financing loan contract and corporate decision-making also influence financing effectiveness, determining the expected profit of both parties with the optimal decision based on game model calculation [14]. Additionally, under the Consignment model, product greenness affects the financing strategy of carbon-dependent manufacturers [15].

Regarding financing risks, the rapid expansion of the e-commerce market has intensified risk issues, with rising uncertainties and risks in e-commerce supply chain bank financing. Therefore, risk identification methods should be actively explored. Scholars have applied Deep Convolutional Neural Networks (DCNN) to predict risks in financing decisions, and DCNN has been shown to provide better predictions than traditional neural networks and support vector machines, with an accuracy rate of 95.36% [16]. Some researchers have utilized game theory to explore the risks in financing decisions and found that higher expected output coefficients result in lower production inputs but increase the purchase price, thus mitigating the impact of risks on financing strategies [17]. Introducing core enterprises into bank financing can leverage their advantageous position and provide guarantees to reduce the overall credit risk in e-commerce supply chains. Game theory analysis, based on the Consignment model, has been used to study financing strategies considering credit guarantees and default risks [18]. Case study research has also shown that risks can be reduced from the perspectives of banking, business, and legal aspects, helping enterprises to take timely measures to mitigate risks.

In terms of financing strategy models, some scholars have constructed game models under the Consignment model to explore the optimal financing strategies for SMEs facing financial constraints [19]. Other researchers have validated financing mechanisms using a cash-to-cash flow sensitivity model [20] and assessed the factors influencing supply chain performance through a comprehensive evaluation method [1]. Additionally, an interpretive structural model has been used to simulate the relationships between financing factors [21]. Bank financing can improve financing efficiency by alleviating information asymmetry. Other researchers have proposed a new cooperation method using a network perspective to trigger capital flows between customers, which better alleviates funding constraints [22]. The integration of artificial intelligence (AI) into e-commerce supply chain bank financing has promising research prospects, with some scholars examining optimal financing strategies using emerging technologies. Under the Direct Selling model, when the regret cost are lower, but the platform's commission is higher, SMEs facing financial constraints are more likely to choose bank financing [23]. Evolutionary game models have been used to study green financing firms, promoting green innovation to achieve balanced finance [24]. Some scholars use Stackelberg game to explore the financially constrained retailers by considering the impact of joint pricing and quality level [25]. When manufacturers and logistics providers are free riders in financing decisions, the game analysis method can be used to obtain the financing equilibrium point and find the optimal financing strategy [26]. In the low-carbon supply chain, considering the impact of low-carbon performance, Stackelberg game is used to explore the optimal financing strategy [27].

Regarding financing strategy optimization, guarantee-based financing solutions have become increasingly popular as they help enterprises more easily access financing and optimize financing strategies. In the Consignment model, in addition to bank financing, alternative finance products and services such as retailer credit finance and hybrid finance are used to offer SMEs more diverse financing options, optimizing the financing strategies [28]. Because the loan limits provided by commercial banks are relatively high, financing costs are also higher, and credit combination financing can broaden the funding channels for constrained enterprises to optimize their financing

strategies [29]. In uncertain conditions, creating value from the perspective of sustainable development can improve corporate credibility and optimize financing strategies [30]. Strengthening green innovation helps attract more consumers and improves financing efficiency [31].

Overall, e-commerce supply chain bank financing has a promising research outlook, bringing more opportunities and advantages for alleviating financing constraints, and can effectively improve financing efficiency. Existing research on the factors influencing bank financing strategies in e-commerce supply chains has paid limited attention to the impact of emerging technologies. This study innovatively explores the effect of blockchain trust mechanisms on bank financing strategies. While existing studies have considered the issue of financing risk when formulating bank financing strategies, there has been little research on how risk attitudes affect financing strategies. Therefore, it is necessary to incorporate the risk aversion attitude of firms into the analysis. Based on social preference theory, this paper introduces the behavioral management of business decision-makers and examines the impact of risk aversion on bank financing strategies, grounded in the real operational scenarios of e-commerce supply chains such as Direct Selling and Consignment. Among various bank financing models, this study mainly adopts the Stackelberg game analysis method, which is different from previous studies in that this study considers the risk aversion and green preference of enterprises and other influencing factors to explore the optimal bank financing strategy. Regarding financing strategy optimization, most existing literature on bank financing strategies has been based on traditional supply chain structures. This paper innovatively investigates optimal financing strategies from the perspective of sales channel choice, closely aligning with the characteristics of e-commerce supply chains and consumer green preferences. It comprehensively considers the two sales models, Consignment and Direct Selling, and different supply chain power structures to explore how different power structures and sales methods impact bank financing strategies.

2.2 Blockchain-based supply chain financing strategy

Research on blockchain-based SCF strategies mainly focuses on the role of blockchain, finance influencing factors under blockchain, the effects of blockchain on financing, and the optimization of financing strategies in blockchain contexts.

Regarding the role of blockchain, blockchain, as a decentralized and secure technology, has injected new vitality and possibilities into SCF, significantly impacting the formulation of financing strategies. Smart contracts can be automatically executed, strengthening trust between enterprises, while consensus mechanisms allow for information sharing and enhance traceability [32]. The coupling of blockchain with SCF enhances the effectiveness of financing strategies within supply chains, breaking through the bottlenecks of traditional financing mechanisms, improving information exchange between enterprises, and assisting banks in better assessing credit levels, thereby providing financing support to financially constrained companies [33]. Using blockchain's smart contracts and other technological tools, the automation of contracts and funds flow, as well as real-time detection, simplify financing processes, reduce human intervention, and mitigate moral hazards such as adverse selection. Machine learning algorithms, applied in blockchain, improve the prediction of corporate credit risks, better preventing risks faced by enterprises and enhancing the security and transparency of blockchain transaction systems [34]. The advantages of blockchain effectively address issues such as difficulty in information sharing and lack of trust mechanisms in the financing process, improving financing efficiency. Through its decentralized and immutable properties, blockchain allows all participants to achieve transparent information sharing and real-time updates, reducing credit risk and increasing the reliability of financing information [35].

Regarding the influencing factors of financing under blockchain, blockchain-based SCF has attracted significant attention. Game-theoretical models are constructed with different blockchain structures to analyze the financing equilibrium and decision-making problems of the participants. It is found that when the number of receivables exceeds a threshold, it becomes more difficult for enterprises to repay loans, and delayed payment financing strategies can help enterprises achieve higher financing benefits. When the loss from financing defaults is relatively low, the likelihood of defaults by financially constrained companies is higher. However, when the default loss is large and reliable income is more substantial, enterprises are more likely to fulfill their commitments, minimizing credit risk and optimizing the financing strategy [36]. In multi-period two-level supply chains, considering the impact of bank insurance on financing strategies, scholars have proposed BCT-SCF platforms that can convert receivables into online bills for transfer payments. With the support of bank insurance, blockchain-driven accounts receivable factoring improves the accessibility of financing. This method is widely applied in business settings. Case studies confirm that the introduction of blockchain technology, higher product pricing, and lower factoring fees all help alleviate financing constraints [37]. A certain degree of product greenness may increase manufacturing costs but boosts overall returns and increases cash flow, thereby alleviating financial constraints. When blockchain investment efficiency is high, trade credit finance is more advantageous than bank finance in improving finance efficiency [38]. The introduction of blockchain enables better tracing of green information, driving green investment, and improving financing availability. Thus, green innovation and investment also influence financing strategies [34]. In the sustainable green supply chain, green investment has attracted much attention. Although the introduction of blockchain technology will increase business costs, it will also increase the level of green sensitivity of consumers and reduce tax costs [39]. Applying blockchain technology to the accounts receivable financing model of green supply chain, Stackelberg was used to explore the optimal financing conditions, and the setting of pledge rate and loan interest rate had a great impact on financing strategies [24]. In the dual-channel green supply chain, financially constrained manufacturers can choose two combination financing strategies: bank financing and platform financing (EB) and prepayment financing and platform financing (EP). Before and after the adoption of blockchain technology, EP is the optimal strategy. When the green degree difference of dual-channel products is large, the offline sales

price is higher than the online sales price. When blockchain costs reach a certain threshold, adoption of blockchain technology should be reduced [40].

In terms of the effects of blockchain on financing, blockchain technology addresses frequent issues such as fraud and high financing costs. It simplifies SCF transaction process and offers solutions to collaborative financing dilemmas. Blockchain can effectively reduce financing risks, and higher blockchain technology and credit costs help improve overall financing returns [41]. Blockchain technology can reshape financing processes, and the new technical architecture based on blockchain's underlying design can resolve trust and default issues within supply chains, perfecting the credit breakdown mechanism, reducing risks, and improving overall efficiency. By providing trustworthy data records and transparent transaction processes, blockchain offers a more reliable and efficient solution to risk reduction [42]. Blockchain helps to strengthen the trust among enterprises in multi-level supply chain. Based on the intuitionistic fuzzy set theory, it implements a reward and punishment mechanism for enterprises through dynamic weight allocation to achieve trust aggregation and credit transmission [43]. The comprehensive trust value of the enterprise can be calculated by using the cloud function in the blockchain, stored in the blockchain and can be split, and the transfer of trust can be realized by the split warrant [44]. In order to alleviate the trust crisis, blockchain technology is used to realize credit information sharing, and a credit evaluation model is built in combination with Funk-SVD to improve the trust framework of supply chain financing [45]. Blockchain's smart contracts have excellent applications in the flow of information and capital operations, better meeting enterprises' funding needs, enhancing financing system transparency and trust, and effectively reducing fraud risks while extending payment deadlines. Blockchain's advanced technologies should be fully leveraged in SCF risk management to improve the effectiveness and feasibility of financing strategies [46]. Financing risks are a major focus for many scholars. In deep-tiered supply chains, upstream suppliers often have small sizes and high financing demands. Limited visibility in multi-tiered supply chains exacerbates financing difficulties. Game analysis shows that blockchain enhances the traceability of on-chain information, optimizes financing decision-making, and mitigates risk-taking by companies. When secondary suppliers face severe financial constraints and their daily operating costs fall below a certain threshold, financing can ensure the interests of tertiary suppliers with lower risks. Cross-level financing based on blockchain can achieve a win-win, reducing financing risks and improving financing efficiency [29].

Regarding the optimization of finance strategies under blockchain, blockchain-powered supply chain financing is widely applied across various industries, significantly promoting innovations in SCF strategies. Some scholars have explored the stable financing equilibrium strategies of supply chain enterprises through a three-party evolutionary game model, analyzing collusion behaviors in enterprises within bank financing, and innovating financing structures and business logic to optimize financing strategies [47]. Other scholars, using case studies of blockchain financing platforms, propose optimization suggestions for financing strategies under blockchain-driven SCF models. By utilizing blockchain's immutability, consensus mechanisms, smart contracts, and other features, transaction processes are effectively simplified, information sharing and transparency are enhanced, trust costs are reduced, and financing difficulties faced by SMEs are alleviated [48]. Blockchain financing platforms accelerate capital flow and integrate supply chain resources based on trustworthy data, providing better solutions to ease enterprises' financing constraints and optimize SCF. As the underlying technology of digital currencies such as Bitcoin, blockchain optimizes financing models and improves liquidity among enterprises on the blockchain. After the introduction of blockchain, issues of distrust between enterprises are resolved, bank loan risks are reduced, the credibility of financing increases, and financing efficiency is significantly improved. To protect corporate privacy data during the financing process, homomorphic encryption can be used to better meet the needs of financing entities, helping financially constrained enterprises secure funding more effectively [49]. In blockchain-enabled SCF, financial institutions' and enterprises' multiple decision-making factors should be comprehensively considered to explore stable equilibrium strategies for financing platforms, solve supply chain distrust and risk issues, assist companies in better managing risks, and optimizing financing strategies [50].

Blockchain technology has a good application in many industries and has an important impact on bank financing strategies. From the existing research, in terms of the role of blockchain, most of them explore its application in financing strategy based on the traditional characteristics of blockchain, such as smart contract. From the perspective of blockchain trust, there are few studies on improving blockchain trust mechanism to optimize financing strategy, while higher trust between e-commerce supply chain and banks can better optimize financing strategy. From the perspective of blockchain trust, this paper explores the influence of blockchain trust mechanism between supply chain and banks, and between supply chain and green consumers on financing strategy. In the aspect of financing influencing factors, although existing scholars have considered risk factors, it is rare to explore its influence on financing strategy from the aspect of risk aversion behavior of enterprises. In terms of financing impact, blockchain can simplify the financing process, reduce the problem of default, better alleviate financing risks and improve financing efficiency. On this basis, this paper introduces contract coordination mechanisms to coordinate interest conflicts and reduce credit risk. Blockchain can improve the trust between enterprises and realize the splitting and transmission of credit. The existing research considers the application of blockchain in green supply chain financing; based on this, this study considers the green preference of consumers and the green investment of suppliers to explore the optimal bank financing strategy. In terms of strategy optimization, most scholars discuss the influence of blockchain on financing strategy based on theory, and some scholars analyze blockchain financing platforms in the form of cases. In this paper, quantitative analysis is used to explore the optimal financing strategy, and at the same time, the influence of risk aversion and sales mode selection on financing strategy is considered to explore how to coordinate the supply chain to optimize financing strategy.

3 Problem description and research hypothesis

3.1 Problem description and notation definition

In an e-commerce green supply chain consisting of suppliers, manufacturers, and green e-commerce platforms, enterprises are risk-averse, two sales methods are considered: Direct Selling and Consignment. These two sales methods are the most common and representative, with significant differences in inventory management, risk taking, capital liquidity and decision-making right allocation, which can cover the typical operation scenarios of different subjects in the supply chain. Under the Direct Selling Mode, suppliers process green products and sell them to market consumers via the e-commerce platform at a price p_s. In this scenario, the supplier must pay the platform a certain usage fee proportion x. With greater pricing power and market control, e-commerce platforms can bring products to market more quickly and obtain cash flow. This model is suitable for those enterprises that pursue rapid financing and want to use the advantages of the platform to reduce operational risks, providing direct profit opportunities for the platform, while also providing a stable sales channel for suppliers. Under the Consignment Mode, suppliers sell green products to manufacturers at a wholesale price w_s. Manufacturers further process these products and sell them to the e-commerce platform at a wholesale price w_d, which then sells them to consumers at a retail price p_e. This approach achieves lower inventory risk and higher price transmission efficiency in the supply chain, which is suitable for those enterprises with low risk but with tight cash flow and helps to improve the financing environment and ease capital constraints. Consumers tend to prefer green, low-carbon products with a higher degree of information transparency. Products with good traceability and environmentally friendly features are more trusted by consumers, and blockchain significantly enhances product traceability and trust. In the daily operation process, risk-averse suppliers need a lot of money to buy raw materials, but their own funds are limited, and they are facing serious financing constraints. Suppliers with tight funds will choose to apply for bank financing.

In traditional bank financing models, banks provide loans to financially constrained suppliers based on guarantees from the e-commerce platform, the financing process is shown in Figure 1a. Most suppliers are SMEs with high credit risk and a lack of trust mechanisms, leading to limited loan amounts from banks. Consequently, risk-averse suppliers rely on the e-commerce platform for credit guarantees, paying a guaranteed fee rate ε. They apply for loans from banks using orders as collateral based on their transactions with the green e-commerce platform. During the loan qualification review process, e-commerce supply chain companies bear the information verification cost. Upon approval, the bank provides loans at an interest rate τ. The supplier uses the funds to process and produce green products, sells them to the green e-commerce platform, and receives accounts receivable. These receivables are then transferred to the bank for factoring. The green e-commerce platform, being risk-averse, repays the bank after the product sales generate returns, charging factoring fees in the interim. The loan interest rate is considered an exogenous variable in this study, which explores the optimal bank financing strategy. In the process of financing decision-making, banks mainly consider their own credit risk, loan cost and other factors. This paper sets the bank loan interest rate as an exogenous variable to explore the bank financing strategy of the e-commerce supply chain.

Under the blockchain trust mechanism, the bank financing process differs, the financing process is shown in Figure 1b. While blockchain adoption incurs additional financing costs, suppliers must pay these costs under both Consignment and Direct Selling Modes. Blockchain improves the authenticity and reliability of information, reducing information verification costs and increasing the bank's trust in e-commerce supply chain enterprises. The industrial and commercial bank of China to launch "ICBC e letter" platform, through intelligent automatically perform contract, reduces the financing cost of upstream suppliers, increasing the utilization of funds. Smart contracts automate loan execution, enabling banks to provide loans to suppliers without guarantees and foster trust transfer. Consensus mechanisms effectively regulate funds, mitigate credit risk during financing, and better accommodate firms' risk-averse preferences. Tencent Cloud combines blockchain technology with financing scenarios. Through multi-party rights confirmation and coordination, the platform helps financial institutions monitor financing risks and effectively solve the financing problems of small and medium-sized enterprises. At the start of the sales period, suppliers sign sales contracts with the green e-commerce and apply for finance from banks using the orders obtained. Banks leverage blockchain to assess the credit information, verify transaction authenticity, repayment capacity, and creditworthiness. Blockchain enhances suppliers' credibility, enabling banks to grant loans. Suppliers then use the funds to process green products and sell them through green e-commerce. After product sales generate returns, funds are promptly allocated to repay the bank. Both members and bank are risk-averse, aiming to maximize supply chain profits while minimizing credit risk. Thus, we explore the impact of blockchain on bank finance in e-commerce green supply chains, considering firms' risk aversion. The bank financing process before and after adopting blockchain is shown in Figure 1.

Based on this, this study constructs a Stackelberg game to explore the optimal bank financing strategy under different sales methods and power structures. Because the e-commerce supply chain usually presents a clear hierarchical structure, enterprises play different roles in the entire process, and each has different decision-making rights and information advantages. Stackelberg game model is based on the "leader-follower" relationship, which can clearly depict the sequential decision-making process of each body in the supply chain. In the e-commerce supply chain composed of suppliers, manufacturers and e-commerce platform, the game analysis method is used to construct the bank financing strategy model under two sales methods of Consignment and Direct Selling. This paper explores the optimal bank financing strategy of e-commerce supply chain enterprises with risk avoidance behavior and consumers with green preference behavior. The two sales methods have different power structures. In Consignment, because the supplier first determines the wholesale price to be sold to the manufacturer, the e-commerce platform then determines the retail price to be sold to the consumer. Therefore, the supplier is the leader, the manufacturer is the follower, and the e-commerce platform is the second follower. Under Direct Selling, the e-commerce platform determines the platform service price, and the supplier determines the selling price on this basis. Therefore, the e-commerce enterprise is the leader, and

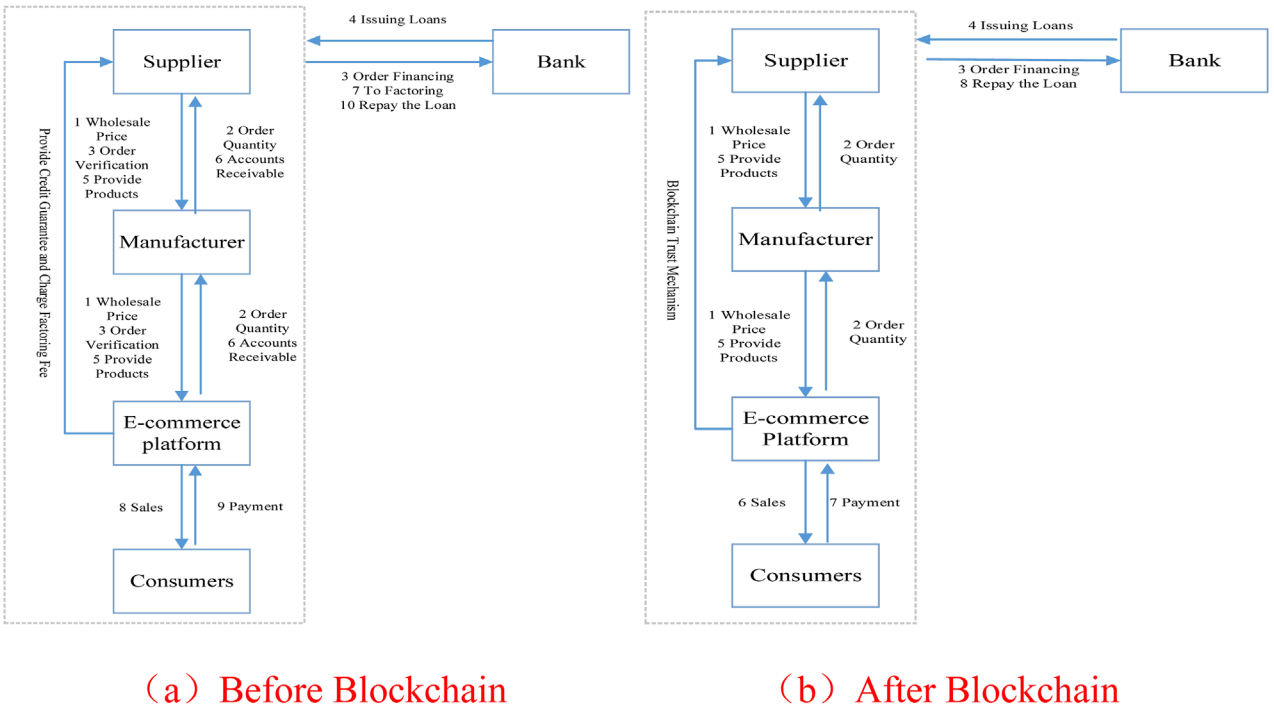


FIGURE 1
Financing processes of E-commerce supply chains before and after blockchain. **(a)** Before Blockchain. **(b)** After Blockchain

the supplier is the follower. Through the Consignment model, we can discuss how to coordinate the game between suppliers as price makers, manufacturers and e-commerce platforms. The Direct Selling model reflects the dominant position of the platform in pricing and financing decisions. This hierarchical structure is exactly consistent with the premise of Stackelberg game model, which provides a solid theoretical foundation for the construction of the model.

In the e-commerce supply chain, in addition to the basic supply and demand relationship, it also involves a variety of factors such as blockchain adoption, risk avoidance behavior, product greenness, consumer green preferences, and guarantee rates. Stackelberg model has high flexibility, which can simultaneously consider multiple influencing factors in the decision-making process of leaders and followers and provide decision support for managers in the face of complex and changeable market environment. The construction model of this study first analyzes the benchmark model of bank financing under Direct Selling and Consignment selling methods and explores the optimal bank financing strategy under supplier financing constraints. Secondly, it analyzes the optimal financing strategy of e-commerce supply chain when enterprises have risk avoidance behavior. Introduce the enterprise's attitude towards risk, explore the optimal bank financing strategy of e-commerce supply chain under risk avoidance. Finally, the introduction of blockchain technology, numerous advantages of blockchain can optimize bank financing strategy, explore the optimal bank financing strategy of risk avoidance e-commerce supply chain under blockchain.

The definitions of the symbols used in this study are shown in Table 1. D represents the market demand function, related to the number of consumers $\tilde{n}$, sales volume q, sales price p, trust a, information verification effort ξ, green preference coefficient μ, product greenness e, Direct Selling proportion β and product valuation v. The financing income π_i and utility function U of supply chain enterprises are affected by product cost c_i, wholesale price w_i, retail price p_i and platform service rate x. The introduction of blockchain technology requires a certain cost b, while reducing the cost of information verification δ, and the blockchain trust coefficient is γ. In the financing process, the loan interest rate of the bank is τ, the factoring rate of return of the e-commerce platform is ε, and the risk sharing ratio is θ. The risk avoidance coefficient is η_i and the green product input coefficient is k.

3.2 Research hypothesis

Based on the two sales methods, Consignment and Direct Selling, in green e-commerce supply chains, this study considers corporate risk aversion and consumers' green preferences to explore the impact of blockchain on bank finance strategies. The hypotheses are proposed:

1. Referring to Tian et al. [51] and Li et al. [52], the market demand follows a random distribution, and consumers have green products' preference, satisfying $D = \tilde{n}\int_{p-a\xi}^{1} f(v)dv = \tilde{n}(q - p + a\xi + \mu e)$, where $f(v)$ is the density function of product valuation v, and $\tilde{n}$ is the potential consumers number. In both Consignment and Direct Selling sales modes, the sales volume follows a Bernoulli distribution:

TABLE 1 Symbol definitions.

Symbol	Meaning
D	Market Demand
$\tilde{n}$	Potential Number of Consumers
ξ	E-Commerce Platform Information Collection Efforts
c_i	Unit Cost
a	Trust Level
τ	Loan Interest Rate
θ	Risk Sharing Ratio
e	Greenness of the Product
π_i	Benefits of E-Commerce Supply Chain Companies
x	Service Fee Rates of E-Commerce Platforms
β	Sales Ratio of Direct Selling
U	The Utility of The Enterprise
w_i	Wholesale Price
p_i	Retail Price
v	Valuation of Products
ε	Factoring Rate of Return
b_i	Blockchain Adoption Costs
δ	Unit Information Verification Cost
μ	Consumer Green Preference Coefficient
k	Green Cost Input Coefficient
γ	Blockchain Trust Factor
η_i	Risk Aversion Factor
q	Quantity of Product
θ	Probability of Supplier Default

$Q = f(x) = \begin{cases} \beta q \\ (1-\beta)q \end{cases}$. Here, q is the potential market demand scale, β denotes the market using Direct Selling, and $1-\beta$ represents the share using the Consignment method. Thus, in Direct Selling mode, the market demand is $D_d = \tilde{n}\left[\beta q - p^d + a\xi + \mu e + \epsilon\right]$. In the Consignment mode, the market demand is $D_c = \tilde{n}\left[(1-\beta)q - p^c + a\xi + \mu e + \epsilon\right]$, where ϵ epsilon is a random variable representing market uncertainty with a mean of 0 and variance σ.

2. Considering that members are risk-averse, the CVaR risk measurement method, adopted by Choi et al. [53] and Gupta and Ivanov [54], is used. The risk-averse utility function is: $U_i = E(\pi_i) - \eta_i\sqrt{VAR(\pi_i)}$. Where $E(\pi_i)$ denotes the expected profit of enterprise, η_i is the risk aversion coefficient of the supply chain enterprise (a higher value indicates greater risk aversion), and $VAR(\pi_i)$ is the variance of the enterprise's stochastic profit: $VAR(\pi_i) = E[\pi_i - E(\pi_i)]^2$.
3. Referring to Jiang et al. [43] and Liu et al. [55], the blockchain trust cost is expressed as $\gamma a^2/2$, where γ is the blockchain trust coefficient, and a represents the degree of blockchain trust, $0 < a \leq 1$. A higher a indicates a higher degree of trust. For inter-enterprise trust, a is negatively correlated with the effort of information collection ξ: the higher the trust, the lower the verification cost, reducing ξ. For

trust between enterprises and consumers, a is positively correlated with price p: the higher the trust, the more people will buy, creating a seller's market, and p will rise accordingly.

4 Model construction

4.1 Bank financing baseline model

In e-commerce supply chains, there are two sales methods: Direct Selling and Consignment. In Direct Selling mode, green e-commerce acts as a service provider and charges platform service fees during sales. In Consignment, the platform is a retailer, purchasing goods from manufacturers and selling to consumers. During the supply chain operation, upstream suppliers face financial constraints and apply for bank financing with guarantees provided by Green E-commerce. Optimal finance strategies under both sales methods are explored.

4.1.1 Direct selling mode

In the Direct Selling mode, suppliers produce products at cost c_s to meet market demand and sell them at price p_s to consumers through green e-commerce. The service fee charged by the platform is x. Suppliers with insufficient initial funds opt for bank financing to alleviate financial constraints. The platform provides a guarantee at a rate ε, and the bank's loan rate is τ.

For suppliers, the inflow of funds is the total revenue from product sales minus the service fee paid to the platform. Outflows include production costs, guarantee fees paid to the platform, and interest expenses paid to the bank. The financing utility function of the supplier is:

$$\max_{p_s^{TD}} \pi_s^{TD} = \left(p_s^{TD} - x^{TD}p_s^{TD}\right)D^{TD} - c_s^{TD}D^{TD}\left(1+\varepsilon^{TD}+\tau^{TD}\right) \tag{1}$$

For the e-commerce platform, inflows are platform service fees and factoring fees from financing, while outflows are platform operating costs. The platform's financing utility function is:

$$\max_{x^{TD},\varepsilon^{TD}} \pi_e^{TD} = x^{TD}p_s^{TD}D^{TD} + c_s^{TD}D^{TD}\varepsilon^{TD} - c_e^{TD}D^{TD} \tag{2}$$

For banks, the overall financing income is mainly interesting income, and the expected income function is as follows:

$$max\tau^{TD}\pi_b^{TD} = c_s^{TD}D^{TD}\tau^{TD} \tag{3}$$

Here, $D^{TD} = \beta q - p^d + \epsilon$.

Proposition 1: According to Equations 1–3, it can be calculated that, the optimal retail price for a financially constrained supplier is: $p_s^{TD^*} = \frac{c_s(1+\tau+\varepsilon)-\beta q(x-1)}{2(1-x)}$, The platform's optimal service fee rate is: $x_e^{TD^*} = \frac{c_e-c_s+\varepsilon-3c_s\varepsilon+c_e\varepsilon-c_s\tau+c_e\tau-2c_s\varepsilon^2-2c_s\varepsilon\tau+\beta q}{4\varepsilon+2\tau+2\tau\varepsilon+2\varepsilon^2+2}$, and the optimal guarantee fee rate is $\varepsilon^{TD^*} = \frac{c_s-c_e+c_s\tau+c_sx-\beta q-\beta qx^2+2x\beta q}{c_s(x-2)}$. Substituting each decision variable into the profit function of the supplier and green e-commerce, the optimal financing utility is $\pi_s^{TD^*} = \frac{(1-x)(c_e+c_s+c_s\tau-\beta q)}{4(2-x)^2}$, $\pi_e^{TD^*} = \frac{(c_e+c_s+c_s\tau-\beta q)^2}{4(2-x)}$.

Proof: Using the backward induction method of Stackelberg theory, the optimal solutions for the decision variables are calculated. According to the supplier's profit, the first and second derivatives of π_s^{TD} about p_s^{TD} are calculated. We obtain $\frac{\partial^2\pi_s^{TD}}{\partial\left(p_s^{TD}\right)^2} = -2 < 0$. It is a concave function, so there is an price that enlarges the supplier's profit. Let $\frac{\partial\pi_s^{TD}}{\partial p_s^{TD}} = 0$, and the best price is $p_s^{TD^*} = \frac{c_s(1+\tau+\varepsilon)-\beta q(x-1)}{2(1-x)}$. $p_s^{TD^*}$ is brought into the expected revenue function of green e-commerce, and then the first derivative and the second derivative of π_e^{TD} about x^{TD} and ε^{TD} are obtained, and the Hessian matrix is found to be negative, That is, π_e^{TD} is a concave function about x^{TD} and ε^{TD}, and there are unique x^{TD} and ε^{TD} to make π_e^{TD} get the optimal value. Let $\frac{\partial\pi_e^{TD}}{\partial x^{TD}} = 0$, $\frac{\partial\pi_e^{TD}}{\partial\varepsilon^{TD}} = 0$, The available optimal platform service fee and optimal guarantee rate are respectively $x^{TD^*} = \frac{c_e-c_s+\varepsilon-3c_s\varepsilon+c_e\varepsilon-c_s\tau+c_e\tau-2c_s\varepsilon^2-2c_s\varepsilon\tau+\beta q}{4\varepsilon+2\tau+2\tau\varepsilon+2\varepsilon^2+2}$ and $\varepsilon^{TD^*} = \frac{c_s-c_e+c_s\tau+c_sx-\beta q-\beta qx^2+2x\beta q}{c_s(x-2)}$. By bringing the variables into the financing utility function, the optimal revenue can be obtained as follows: $\pi_s^{TD^*} = \frac{(1-x)(c_e+c_s+c_s\tau-\beta q)}{4(2-x)^2}$, $\pi_e^{TD^*} = \frac{(c_e+c_s+c_s\tau-\beta q)^2}{4(2-x)}$.

Inference 1 The expected profit π_s^{TD} of suppliers will decrease with an increase in market demand βq under the Direct Selling mode. Under the Direct Selling mode, the larger market scale is due to the measures taken by suppliers to reduce prices, and the lower sales price attracts more consumers and speeds up the withdrawal of funds. However, in the larger market demand βq, the platform service fee expenditure increases accordingly, so the overall higher market share is not conducive to the increase of suppliers' income, and it cannot better solve its own financial problems. When the service fee income x^{TD} is large, the revenue level $\pi_e^{TD^*}$ will increase and the cash flow will be large. The guaranteed fee income will decrease with the increase of market demand βq, and the platform service fee income will increase with the market share βq, but the profit level $\pi_e^{TD^*}$ will decrease under higher market demand. The platform service fee should be within a certain limit. Excessive service rate x will damage the enthusiasm of suppliers, but it is not good for improving the finance income of green e-commerce.

4.1.2 Consignment mode

Under the Consignment model, the supplier produces goods at c_s and sells them to the manufacturer at w_s. The manufacturer further processes the products, incurring a cost of c_d, and sells them to the green e-commerce platform at w_d. Finally, the green e-commerce platform

sells to green consumers at p_e. When the supplier faces financial constraints and applies for bank financing, they pay factoring fees at a guaranteed rate ε and loan interest to the bank at a rate τ.

For the supplier, the primary cash inflow is the total revenue from product sales. The cash outflows include production costs, financing guarantee fees, and loan interest payments. The financing utility of supplier is:

$$\max_{w_s^{TC}} \pi_s^{TC} = w_s^{TC} D^{TC} - c_s^{TC} D^{TC}\left(1 + \varepsilon^{TC} + \tau^{TC}\right) \tag{4}$$

For the manufacturer, the main operation involves processing the purchased products into finished goods. The cash inflow is the revenue from selling to the green e-commerce platform, while the cash outflows include the costs of purchasing products and further processing. The financing utility of manufacturers is:

$$\max_{w_d^{TC}} \pi_d^{TC} = \left[w_d^{TC} - w_s^{TC} - c_d^{TC}\right] D^{TC} \tag{5}$$

For green e-commerce, the primary cash inflows are revenues from product sales and guarantee fee income. The main cash outflows are the costs of purchasing products from the manufacturer. The financing utility of green e-commerce is:

$$\max_{p_e^{TC}, \varepsilon^{TC}} \pi_e^{TC} = p_e^{TC} D^{TC} + c_s^{TC} D^{TC} \varepsilon^{TC} - w_d^{TC} D^{TC} \tag{6}$$

For banks, the capital inflow is mainly interesting income, and the expected income function is as follows:

$$max \tau^{TC} \pi_b^{TC} = c_s^{TC} D^{TC} \tau^{TC} \tag{7}$$

Here, $D^{TC} = (1-\beta)q - p^c + \epsilon$

Proposition 2: Under the Consignment method of bank financing, according to Equations 4–7, it can be calculated that, the supplier's optimal wholesale price is $w_s^{TC^*} = \frac{c_s - c_d + 3w_d - c_s\varepsilon + c_s\tau - 2q + 2q\beta}{2}$, the manufacturer's optimal wholesale price is $w_d^{TC^*} = \frac{c_d + w_s + c_s\varepsilon + (1-\beta)q}{2}$, the green e-commerce's optimal retail price is $p_e^{TC^*} = \frac{w_d - c_s\varepsilon + (1-\beta)q}{2}$, and the optimal guarantee rate is $\varepsilon^{TC^*} = \frac{w_d - (1-\beta)q}{c_s}$. The optimal profit function of the capital-constrained supplier is $\pi_s^{TC^*} = \frac{(c_s - c_d - 4q - 2w_s + 5w_d - 3c_s\varepsilon + c_s\tau + 4\beta q)(c_s + 2c_d + c_s\tau + q + 2c_s\varepsilon - \beta q + w_s - 3w_d)}{16}$, the manufacturer's optimal financing utility under the Consignment sales method is $\pi_d^{TC^*} = \frac{(c_d + q + w_s - 2w_d + c_s\varepsilon - \beta q)(c_s - w_s + 3w_d - 2c_s\varepsilon + c_s\tau - 3q + 3\beta q)}{8}$, and the green e-commerce's optimal financing utility is $\pi_e^{TC^*} = \frac{(c_d + w_s - 2w_d + c_s\varepsilon + q - \beta q)^2}{16}$.

Proof: The backward induction method of Stackelberg game is used. First, the first and second derivatives of π_e^{TC} about p_e^{TC} and ε^{TC} are obtained, and the Hessian matrix is calculated, so it can be judged that the matrix is negative, that is to say, π_e^{TC} is a concave function about $\left(p_e^{TC}, \varepsilon^{TC}\right)$, and there are optimal ε^{TC} and p_e^{TC}, which makes π_e^{TC} get the optimal solution. Let $\frac{\partial \pi_e^{TC}}{\partial p_e^{TC}} = 0$, $\frac{\partial \pi_e^{TC}}{\partial \varepsilon^{TC}} = 0$, The optimal solution is $p_e^{TC^*} = \frac{w_d - c_s\varepsilon + (1-\beta)q}{2}$ and $\varepsilon^{TC^*} = \frac{w_d - (1-\beta)q}{c_s}$. The $p_e^{TC^*}$ is brought into the manufacturer's expected revenue function, and the first and second partial derivatives of π_d^{TC} about w_d^{TC} are obtained, $\frac{\partial^2 \pi_d^{TC}}{\partial \left(w_d^{TC}\right)^2} = -1 < 0$. The π_d^{TC} is a concave function, and the existence of the optimal price makes the manufacturer's expected revenue maximum. Let $\frac{\partial \pi_d^{TC}}{\partial w_d^{TC}} = 0$, The optimal wholesale price of available manufacturers is $w_d^{TC^*} = \frac{c_d + w_s + c_s\varepsilon + (1-\beta)q}{2}$. Bring the above decision variables into π_s^{TC}, and solve the first and second partial derivatives of π_s^{TC} about w_s^{TC}, $\frac{\partial^2 \pi_s^{TC}}{\partial \left(w_s^{TC}\right)^2} = -1/2 < 0$. The expected income function of suppliers with financial constraints is concave function, and the optimal solution of wholesale price makes the expected income of supplier's maximum. Let $\frac{\partial \pi_s^{TC}}{\partial w_s^{TC}} = 0$. The optimal price is $w_s^{TC^*} = \frac{c_s - c_d + 3w_d - c_s\varepsilon + c_s\tau - 2q + 2q\beta}{2}$. By bringing each decision variable into the expected revenue function, we get: $\pi_s^{TC^*} = \frac{(c_s - c_d - 4q - 2w_s + 5w_d - 3c_s\varepsilon + c_s\tau + 4\beta q)(c_s + 2c_d + c_s\tau + q + 2c_s\varepsilon - \beta q + w_s - 3w_d)}{16}$, $\pi_d^{TC^*} = \frac{(c_d + q + w_s - 2w_d + c_s\varepsilon - \beta q)(c_s - w_s + 3w_d - 2c_s\varepsilon + c_s\tau - 3q + 3\beta q)}{8}$, $\pi_e^{TC^*} = \frac{(c_d + w_s - 2w_d + c_s\varepsilon + q - \beta q)^2}{16}$.

Inference 2 The profit $\pi_e^{TC^*}$ will increase with the increase of market demand $(1-\beta)q$ of consideration mode. When βq is small, that is, $(1-\beta)q$ is large, the sales price p_e^{TC} will be relatively high, the optimal profit level will be improved, and more financing income will be obtained. Under the higher market demand of consideration mode $(1-\beta)q$, a buyer's market will be formed, and the supplier as the leader will set a lower sales price $w_s^{TC^*}$, which is not conducive to the increase of the supplier's cash flow. However, the sales prices w_d^{TC} and p_e^{TC} of manufacturers and e-commerce platforms will increase correspondingly under higher market demand $(1-\beta)q$, and the revenue level of green e-commerce $\pi_e^{TC^*}$ will increase with the increase of market demand $(1-\beta)q$, thus improving the overall cash flow. The optimal guaranteed rate ε^{TS} will decrease with the increase of market $(1-\beta)q$, and the financing cost of suppliers will decrease under higher market demand. The optimal guaranteed interest rate ε^{TS} also decreases with the increase of the supplier's cost c_s^{TC}, and the e-commerce platform will appropriately decrease the guaranteed fee for the sake of protecting the interests of suppliers.

4.2 Risk-averse E-Commerce supply chain bank financing

Consumers generally have green preferences, and suppliers tend to produce green and low-carbon products when making operational decisions. When suppliers produce green products, the cost of green technology input is $ke^2/2$, which is a one-time technology investment

and will not affect the unit production cost, μ^2/k indicating the green efficiency coefficient. In the process of financing, e-commerce supply chain enterprises are all risk-averse, and it is necessary to pay a certain information verification cost to verify financing information.

4.2.1 Direct selling mode

For the financially constrained supplier, the cash inflow under the Direct Selling model is the total revenue of product sales minus the service fees paid to green e-commerce. The cash outflows include the costs of purchasing and processing raw materials, guarantee fees paid to green e-commerce, interest expenses on bank loans, and costs incurred in providing documentation for financing. The financing utility of supplier is:

$$\max_{p_s^{RD},e^{RD}} \pi_s^{RD} = \left(p_s^{RD} - x^{RD}p_s^{RD}\right)D^{RD} - \left(c_s^{RD}D^{RD} + k^{RD}e^2/2\right)\left(1 + \varepsilon^{RD} + \tau^{RD} + \delta^{RD}\right) \tag{8}$$

For e-commerce platforms, the main services provided are for suppliers to sell products and obtain financing. The primary source of funding inflow is the platform service fees charged for providing a platform to sell products and the factoring fees charged for providing guarantees to suppliers. The resulting cash outflows include the operating costs of green e-commerce and the costs of reviewing loan documentation, among others. In addition, there are risk-sharing contracts between enterprises. If a supplier fails to repay the funds upon maturity, the e-commerce enterprise must repay the funds on their behalf. The probability of supplier defaulting is denoted by θ, so the green e-commerce's financing utility is:

$$\max_{x^{RD},\varepsilon^{RD}} \pi_e^{RD} = x^{RD}p_s^{RD}D^{RD} + \left(c_s^{RD}D^{RD} + k^{RD}e^2/2\right)\varepsilon^{RD} - c_e^{RD}D^{RD} - \left(c_s^{RD}D^{RD} + k^{RD}e^2/2\right)\left[\delta^{RD} + \theta^{RD}\left(1 + \tau^{RD}\right)\right] \tag{9}$$

For banks, capital inflow is an interest income, and the income function is as follows:

$$\max_{\tau^{RD}} \pi_b^{RD} = \left(c_s^{RD}D^{RD} + k^{RD}e^2/2\right)\tau^{RD} \tag{10}$$

Here, $D^{RD} = \beta q - p^d + \mu e + \epsilon$.

Considering that enterprises are risk aversion, the CVaR risk measurement method is used to characterize the risk behavior. Based on the function, the utility functions of suppliers and e-commerce platforms considering risk aversion can be obtained as: $U\left(\pi_s^{RD}\right) = \left(p_s^{RD} - x^{RD}p_s^{RD}\right)\left(\beta q - p^d - \eta\sigma + \mu e\right) - \left[c_s^{RD}\left(\beta q - p^d - \eta\sigma + \mu e\right) + k^{RD}e^2/2\right]\left(1 + \varepsilon^{RD} + \tau^{RD} + \delta^{RD}\right)$ and $U\left(\pi_e^{RD}\right) = x^{RD}p_s^{RD}\left(\beta q - p^d - \eta\sigma + \mu e\right) + \left[c_s^{RD}\left(\beta q - p^d - \eta\sigma + \mu e\right) + ke^2/2\right]\varepsilon^{RD} - c_e^{RD}\left(\beta q - p^d - \eta\sigma + \mu e\right) - \left[c_s^{RD}\left(\beta q - p^d - \eta\sigma + \mu e\right) + k^{RD}e^2/2\right]\left[\delta^{RD} + \theta^{RD}\left(1 + \tau^{RD}\right)\right]$. By solving the first-order derivative and the second-order derivative, it can be concluded that $U\left(\pi_s^{RD}\right)$ and $U\left(\pi_e^{RD}\right)$ are concave functions and there is an optimal solution.

Proposition 3: According to Equations 8–10, it can be calculated that, the optimal price of the supplier is $p_s^{RD^*} = \frac{(\delta+\varepsilon+\tau+1)\begin{pmatrix} c_s k - c_s\mu^2 + c_s\mu^2 x + c_s\delta k + c_s k\varepsilon + ak\xi + c_s k\tau \\ -\sigma k\eta + k\beta q - ak\xi x + k\eta\sigma x - kx\beta q \end{pmatrix}}{(1-x)(2k+2\delta k+2k\varepsilon+2k\tau+\mu^2 x-\mu^2)}$, the optimal product greenness is $e^{RD^*} = \frac{\mu\ (c_s+c_s\ \ \delta+c_s\ \ \varepsilon-a\xi+c_s\ \ \tau+\eta\sigma-\beta q+a\ \ \xi x-\eta\sigma\ \ x+x\ \ \beta q)}{2\ \ k+2\ \ \delta\ \ k+2\ \ k\varepsilon+2\ \ k\tau+\mu^2\ \ x-\mu^2}$, and the optimal platform fee rate is $x^{RD^*} = \frac{c_e - c_s - \eta\sigma + \mu e + \beta q + \delta}{2}$, the optimal guaranteed rate is $\varepsilon^{RD^*} = \frac{\mu^2(c_e+c_s\ \ \theta-p_s\ \ x+c_s\ \ \delta+c_s\ \ \theta\ \ \tau\ )+2\ \ k(\delta p_s+\theta\ \ p_s+\delta\eta\sigma+\eta\sigma\theta-\delta\beta q+\theta p_s\ \ \tau-\theta\beta q+\eta\sigma\theta\tau-\theta\ \ \beta q\tau)}{c_s\ \ \mu^2+2\ \ k\ \ (p_s+\eta\sigma-\beta q)}$

The optimal financing utility of the supplier is $\pi_s^{RD^*} = \frac{k\ \ (\delta+\varepsilon+\tau+1)\ \ (c_s\ \ (\delta+\varepsilon+\tau+1)+\beta q(x-1))^2}{2\ \ (1-x)\ \ (2\ \ k(\delta+\varepsilon+\tau+1)+\mu^2\ \ (x-1))}$, and the optimal financing utility of the e-commerce platform is $\pi_e^{RD^*} = (c_e - xA)(\eta\sigma - \mu e - \beta q + A) - \left(\frac{e^2 k}{2} - c_s(\eta\sigma - \mu e - \beta q + A)\right)(\delta - \varepsilon + \theta(\tau+1))A = \frac{((1-x)(\mu e+\beta q)+c_s(\delta+\varepsilon+\tau+1))}{2(1-x)}$

Proof: Using the backward induction method of Stackelberg, the verification process is the same as Proposition 1.

Inference 3 Under the high market demand βq of the Direct Selling model, the sales price p_s^{RC} is relatively high, which helps to recover funds faster. A higher green preference coefficient k will also prompt suppliers to increase sales prices, but the larger the green efficiency coefficient μ^2/k, the more likely it is that suppliers will set lower sales prices, pass on benefits to consumers, attract more customers, and strive to obtain more profits. The greenness of products e^{PC^*} will increase with the increase of suppliers' production costs c_s^{RD}, and is positively correlated with market demand βq under the Direct Selling model, indicating that under higher market demand, the greener the product, the more consumers will be attracted to recover funds faster. The optimal guaranteed rate ε^{RD} will increase with the increase of the green efficiency μ^2/k, increasing the financing cost of suppliers. Under higher Direct Selling market demand βq, suppliers' profits π_s^{RD} will increase with the increase of market demand βq, which will help to better alleviate capital constraints.

4.2.2 Consignment mode

The main source of funding inflow for the supplier is the sales revenue obtained from selling products to the manufacturer, while the funding outflows include the production cost of raw materials, green investment costs, factoring fees required for guarantees, interest expenses for applying for financing from the bank, and costs for providing information certification. The supplier's financing utility function is:

$$\max_{w_s^{RC},k^{RC}} \pi_s^{RC} = w_s^{RC}D^{RC} - \left(c_s^{RC}D^{RC} + k^{RC}e^2/2\right)\left(1 + \varepsilon^{RC} + \tau^{RC} + \delta^{RC}\right) \tag{11}$$

For the manufacturer, the primary sources of funding inflow mainly include the income obtained from selling products, while the funding outflows incurred include the product procurement cost, production and operation costs, and information verification costs. The financing utility function of the manufacturer is:

$$\max_{w_d^{RC}} \pi_d^{RC} = \left(w_d^{RC} - w_s^{RC} - c_d^{RC}\right)D^{RC} - \left(c_s^{RC}D^{RC} + k^{RC}e^2/2\right)\delta^{RC} \tag{12}$$

For the green e-commerce platform, the sources of funding inflow include the revenue from selling products and the factoring fees obtained. The funding outflows incurred include the costs of purchasing products and the expenses for information review. To coordinate the interest relationships, a risk-sharing contract is introduced. If the supplier does not repay the loan upon maturity, the green e-commerce needs to pay the default amount according to the ratio θ. Therefore, green e-commerce's financing utility is:

$$\max_{p_e^{TC},\varepsilon^{RC}} \pi_e^{RC} = p_e^{RC}D^{RC} + \left(c_s^{RC}D^{RC} + k^{RC}e^2/2\right)\varepsilon^{RC} - w_d^{RC}D^{RC} - \left(c_s^{RC}D^{RC} + k^{RC}e^2/2\right)\left(\delta^{RC} + \theta^{RC}\left(1+\tau^{RC}\right)\right) \tag{13}$$

For banks, capital inflow is the interest earned, and the income function is:

$$\max_{\tau^{RC}} \pi_b^{RC} = \left(c_s^{RC}D^{RC} + k^{RC}e^2/2\right)\tau^{RC} \tag{14}$$

Here, $D^{RC} = (1-\beta)q - p^c + \mu e + \epsilon$.

Since the firms are risk-averse, the CVaR risk measurement method is used to introduce the mean - standard deviation to characterize the risk-averse behavior of suppliers and e-commerce platforms. According to the profit function, the financing utility of the supplier and the green e-commerce are: $U\left(\pi_s^{BC}\right) = w_s^{RC}[(1-\beta)q - p^c + \mu e - \eta\sigma] - \left[c_s^{RC}[(1-\beta)q - p^c + \mu e - \eta\sigma] + k^{RC}e^2/2\right]\left(1+\varepsilon^{RC}+\tau^{RC}+\delta^{RC}\right)$ and $U\left(\pi_e^{BC}\right) = p_e^{RC}[(1-\beta)q - p^c + \mu e - \eta\sigma] + \left[c_s^{RC}[(1-\beta)q - p^c + \mu e - \eta\sigma] + k^{RC}e^2/2\right]\varepsilon^{RC} - w_d^{RC}[(1-\beta)q - p^c + \mu e - \eta\sigma] - \left(c_s^{RC}[(1-\beta)q - p^c + \mu e - \eta\sigma] + k^{RC}e^2/2\right)\left(\delta^{RC} + \theta^{RC}\left(1+\tau^{RC}\right)\right)$. By solving the first-order derivative and the second-order derivative, it can be concluded that $U\left(\pi_s^{RC}\right)$ and $U\left(\pi_e^{RC}\right)$ are concave functions, so there is an optimal solution.

Proposition 4: According to Equations 11–14, it can be calculated that, the optimal retail price is $p_e^{RC^*} = \frac{k\ e^2+2\ c_s\mu\ e+2\ c_s\ q-2\ c_s\ \eta\sigma-2\ c_s\ q\beta}{2\ c_s}$ and the optimal guaranteed rate is $\varepsilon^{RC^*} = \frac{c_s\ w_d-c_s\ q+{c_s}^2\ \delta-e^2\ k+{c_s}^2\theta+{c_s}^2\ \theta\tau+c_s\ \eta\sigma-c_s\ e\ \mu+c_s\ q\beta}{{c_s}^2}$. The optimal price of the manufacturer is ${w_d}^{RC^*} = \frac{c_d+w_s+c_s\ \delta-\sigma\ n+e\ \mu-q\ (\beta-1)-c_s\ (\delta-\varepsilon+\theta\ (\tau+1))}{2}$. The optimal price of the supplier is ${w_s}^{RC^*} = \frac{c_s-c_d-2q+3\ w_d+2c_s\ \delta-c_s\ \varepsilon+2c_s\ \theta+c_s\ \tau+2\sigma\ n-2e\ \mu+2q\ \beta-3\ e^2\ k+2c_s\ \theta\ \tau}{2}$ and the optimal product green preference coefficient is $k^{RC^*} = \frac{c_s\ (c_s+c_d-q+2\ w_s+3\ c_s\ \delta+c_s\ \theta+c_s\ \tau+\sigma\eta-e\ \mu+q\ \beta+c_s\ \theta\ \tau)}{2\ e^2}$. The optimal financing utility of the supplier is $\pi_s^{RC^*} = \frac{{c_s}^2(9\ \ \delta^2+2(\theta\tau+3\ \ \delta+1)(\theta\tau+\theta+\tau+1)-\theta^2\ \ \tau^2+\theta^2+\tau^2-1)+(2c_s(\theta\tau+\theta+\tau+1+3\delta)+2c_d)(c_d-e\ \ \mu+\eta\sigma-q+q\ \ \beta)-{c_d}^2}{16}$, $+\frac{c_s e^2 k(\ e\ \ \mu-\eta\sigma-q\beta+q-w_d\ \ )+e^4\ \ k^2}{2\ {c_s}^2}$

the optimal financing utility of the manufacturer is $\pi_d^{RC^*} = c_d\left(\frac{c_d-q+w_s+2c_s\ \ \delta-c_s\ \ \varepsilon+c_s\ \ \theta+\eta\sigma-e\mu+q\beta+c_s\ \ \theta\tau}{4}\right)+c_s\delta\ \ \left(\frac{c_d+w_s-w_d}{4}\right)$, and the optimal financing utility of the green e-commerce is $\pi_e^{RC^*} = \frac{(c_d+w_s+c_s\ \ \delta+\eta\sigma-e\ \ \mu+q\ \ (\beta-1)+c_s\ \ (\delta-\varepsilon+\theta\ \ (\tau+1)))^2}{16}$ $-\frac{(2e^2\ \ k-c_s(c_d+w_s+c_s\ \ \delta+\eta\sigma-e\ \ \mu+q\ \ (\beta-1)+c_s\ \ (\delta-\varepsilon+\theta(\tau+1))))(\delta-\varepsilon+\theta(\tau+1))}{4}$.

Proof: According to the backward induction method of Stackelberg game, the calculation process of the optimal solution of variables is like Theorem 2, so I will not repeat it here.

Inference 4 The optimal selling price p_e^{RC} will increase with the increase of the green degree. In the supplier-led Consignment mode, under the higher Consignment market $(1-\beta)q$, the price of the green e-commerce platform p_e^{RC} will increase, which will help to obtain more funds, which is completely opposite Direct Selling. The price p_e^{RC} is negatively correlated with the risk aversion η^{RC}. As the risk aversion η^{RC} increases, the selling price will be reduced to attract more consumers to obtain funds. The optimal guarantee rate ε^{RC} will increase with the increase of information verification efforts δ^{RC} , but will decrease with the increase of product green k^{RC} and product green degree e^{RC} , and will also decrease with the increase of Consignment sales market demand $(1-\beta)q$, and will also increase with the increase of supplier default probability θ. The manufacturer's best-selling price w_d^{RC} will increase with the increase in Consignment sales market size $(1-\beta)q$ and is negatively correlated with the default probability θ. The supplier's selling price w_s^{RS} will rise with the default probability θ and will rise with the increase of Consignment sales market size $(1-\beta)q$.

4.3 Risk-averse e-commerce supply chain bank financing under blockchain

Under the blockchain, enterprises upload transaction and financing information to the chain in time, which saves the cost of information verification, ensures the authenticity and reliability of information, and effectively reduces financing risks. The introduction of blockchain has increased the technical cost b, promoted the establishment of trust mechanisms, and strengthened the trust between enterprises and between enterprises and consumers. The trust cost paid is $\gamma a^2/2$.

4.3.1 Direct selling mode

Under blockchain technology, financially constrained suppliers need to adopt blockchain, which increases blockchain-related costs. Cash inflows primarily consist of the total revenue from product sales minus the service fees paid to the green e-commerce platform. Cash outflows include the costs of purchasing and processing materials, the costs of blockchain technology, green investment costs, and a blockchain trust cost of $\gamma a^2/2$, which helps facilitate fund auditing and product traceability. Additionally, suppliers incur loan interest expenses. The financing utility function of supplier is:

$$\max_{p_s^{BD},e^{BD}} \pi_s^{BD} = \left(p_s^{BD} - x^{BD}p_s^{BD}\right)D^{BD} - \left[\left(c_s^{BD} + b_s^{BD}\right)D^{BD} + k^{BD}e^2/2 + \gamma^{BD}a^2/2\right]\left(1 + \tau^{BD}\right) \tag{15}$$

An e-commerce platform that adopts blockchain technology needs to incur certain technical introduction costs to enhance the reliability of financing information. Suppliers selling products on the platform can obtain product sales service fees. The cash outflows include the introduction costs of blockchain, the operating costs of the platform, and the blockchain trust costs. Therefore, the financing utility of the green e-commerce platform is:

$$\max_{x^{BD},a^{BD}} \pi_e^{BD} = \left(x^{BD}p_s^{BD} - c_e^{BD} - b_e^{BD}\right)D^{BD} - \gamma^{BD}a^2/2 \tag{16}$$

For banks, capital inflow is mainly interest earned, and the income function is as follows:

$$\max_{\tau^{BD}} \pi_b^{BD} = \left[\left(c_s^{BD} + b_s^{BD}\right)D^{BD} + k^{BD}e^2/2 + \gamma^{BD}a^2/2\right]\tau^{BD} \tag{17}$$

Here, $D^{BD} = \beta q - p^d + a\xi + \mu e + \epsilon$.

Since all enterprises are risk-averse, the mean–standard deviation method is used to characterize the risk-averse, that is $U\left(\pi_i^{BD}\right) = E\left(\pi_i^{BD}\right) - \eta_i\sqrt{var\left(\pi_i^{BD}\right)}$, the financing utility of suppliers and green e-commerce platforms are: $U\left(\pi_s^{BD}\right) = \left(p_s^{BD} - x^{BD}p_s^{BD}\right)\left[\beta q - p^d + a\xi + \mu e - \eta\sigma\right] - \left[\left(c_s^{BD} + b_s^{BD}\right)\left[\beta q - p^d + a\xi + \mu e - \eta\sigma\right] + k^{BD}e^2/2 + \gamma^{BD}a^2/2\right]\left(1 + \tau^{BD}\right)$ and $U\left(\pi_e^{BD}\right) = \left(x^{BD}p_s^{BD} - c_e^{BD} - b_e^{BD}\right)\left[\beta q - p^d + a\xi + \mu e - \eta\sigma\right] - \gamma^{BD}a^2/2$. It is easy to prove that $U\left(\pi_s^{BD}\right)$ and $U\left(\pi_e^{BD}\right)$ are concave functions, so there is an optimal solution.

Proposition 5: Under the e-commerce supply chain bank financing model driven by blockchain technology, according to Equations 15–17, it can be calculated that, the supplier's optimal retail price is $p_s^{BD^*} = \frac{(\tau+1)\quad(k(b+c_s+a\xi+b\quad\tau+c_s\tau-\eta\sigma+q\beta-a\xi x+\eta\sigma\quad x-q\quad x\beta)-\mu^2(b+c_s-b\quad x-c_s\quad x))}{(x-1)\quad(2\quad k+2\quad k\quad\tau+\mu^2\quad x-\mu^2)}$, the product's optimal green preference is $e^{BD^*} = \frac{\mu\quad(b_s+c_s-a\quad\xi+b_s\quad\tau+c_s\tau+\sigma\eta-q\beta+a\quad\xi x-\sigma\quad\eta x+q\quad x\beta)}{2\quad k+2\quad k\tau+\mu^2\quad x-\mu^{2\prime}}$, and the optimal trust is $a^{BD^*} = \frac{\xi\begin{pmatrix} p\mu^2\quad x(x-1)\quad(2\quad k+2\quad k\quad\tau+\mu^2\quad x-\mu^2)+(2\quad b\quad k^2+2\quad c_e\quad k^2+2\quad\eta\sigma\quad k^2x)(\tau+1)^2 \\ +(b\mu^4+c_e\quad\mu^4)(1-x)^2+(3b\quad k\quad\mu^2+3c_e\quad k\quad\mu^2+\eta\sigma k\quad\mu^2\quad x+k\quad q\mu^2\quad x\beta+e\quad k\mu^3\quad x)(\tau+1)(1-x) \\ -2\quad k^2\quad q\quad x\beta\tau(2+\tau\quad)+a\quad k\quad\xi\quad\mu^2x(x-\tau-1)-2\quad e\quad k^2\quad\mu x(1-\tau)^2+c_s\quad k\quad\mu^2\quad x \end{pmatrix}}{\gamma(2\quad k+2\quad k\quad\tau+\mu^2\quad x-\mu^2)^2-(2\quad k\quad\xi^2\quad x\quad(\tau+1)\quad(k+k\tau+\mu^2\quad x-\mu^2))}$. The green e-commerce platform's optimal service fee satisfies $x_p^{BD^*} = \frac{c_e-c_s+a\xi-\eta\sigma+\mu e+\beta q+b_e^{BC}}{2}$. The supplier's optimal financing utility is $\pi_s^{BD^*} = \frac{p\begin{pmatrix} p(x-1)\quad(2\quad k+2\quad k\tau+\mu^2\quad x-\mu^2)/(\tau+1)+(ek\tau(2\quad k+2\quad k\tau+\mu^2\quad x-\mu^2)/\mu)+2kx(a\quad\xi-\eta\sigma+q\quad\beta) \\ +\tau\mu^2(x-1)(b+c_s\quad)+((\mu^2-2k-2k\tau-\mu^2\quad x)+(u^2\quad x^2+2\quad k\quad\tau\quad x-\mu^2x))(a\xi-\eta\sigma+q\beta+e\mu) \end{pmatrix}}{(2\quad k+2\quad k\tau+\mu^2\quad x-\mu^2)}$, and the e-commerce platform's optimal financing utility is $\pi_e^{BD^*} = \frac{a^2\quad\gamma}{2} - \frac{(b+c_e+xp)\begin{pmatrix} p(x-1)\quad(2\quad k+2\quad k\quad\tau+\mu^2\quad x-\mu^2)+(2\tau+1) \\ (\quad k\quad\eta\sigma-k\quad q\quad\beta+a\quad k\quad\xi\quad x-\eta\sigma\quad k\quad x+k\quad q\quad x\quad\beta) \\ +(\mu^2+\mu^2\quad x^2-2\quad\mu^2\quad x)(e\mu+a\quad\xi-\eta\sigma+q\beta) \\ +2k(\quad e\quad\tau\quad\mu\quad x-a\quad\xi-e\quad\mu-a\quad\xi\tau-e\quad\tau\mu+e\quad\mu\quad x) \end{pmatrix}}{(x-1)\quad(2\quad k+2\quad k\tau+\mu^2\quad x-\mu^2)}$.

Proof: The solution is obtained by the reverse induction method of Stackelberg game. The calculation process is the same as the derivation process of Theorem 1 and will not be repeated here.

Inference 5 The supplier's optimal retail price p_s^{BD} is positively correlated with blockchain technology b_s^{BD} and product production costs c_s^{BD}, and will increase with the loan interest rates τ, obtaining more capital flows. The higher the service fee x, the higher the retail price p_s^{BD} will be, because the supplier's cost increases. Higher blockchain trust a^{BD} will lead to higher retail prices p_s^{BD}, which will help to recover funds faster. Retail prices p_s^{BD} will decrease with the increase of risk aversion η^{BD}, but are positively correlated with the market size βq and the product green preference coefficient k^{BD}. The optimal greenness of the product is negatively correlated with the size of the market size βq, and will also decrease with the increase of blockchain trust a^{BD}. Blockchain trust is positively correlated with the risk aversion η^{BD} and blockchain technology b^{BD}, but will decrease with the increase of market size βq.

4.3.2 Consignment mode

For the supplier, cash inflows primarily consist of the total revenue from product sales. Cash outflows include the cost of procuring raw materials, the cost of introducing blockchain technology, green investment costs, and blockchain trust costs. Additionally, suppliers must pay bank loan interest expenses. The supplier's financing utility function is:

$$\max_{w_s^{BC},e^{BC}} \pi_s^{BC} = w_s^{BC} D^{BC} - \left[\left(c_s^{BC} + b_s^{BC}\right) D^{BC} + k^{BC} e^2/2 + \gamma^{BC} a^2/2\right]\left(1 + \tau^{BC}\right) \tag{18}$$

For the manufacturer, cash inflows are mainly derived from revenue generated by selling products. Cash outflows include the cost of purchasing and processing products, the cost of adopting blockchain technology, and blockchain trust costs. The manufacturer's financing utility function is:

$$\max_{w_d^{BC}} \pi_d^{BC} = \left(w_d^{BC} - w_s^{BC} - c_d^{BC} - b_d^{BC}\right) D^{BC} - \gamma^{BC} a^2/2 \tag{19}$$

In the process of SCF, the green e-commerce platform generates cash inflows from the revenue earned by selling products to consumers. Cash outflows include the cost of procuring products, the cost of introducing blockchain technology, and blockchain trust costs. Green e-commerce's financing utility function is:

$$\max_{p_e^{BC},a^{BC}} \pi_e^{BC} = \left(p_e^{BC} - w_d^{BC} - b_e^{BC}\right) D^{BC} - \gamma^{BC} a^2/2 \tag{20}$$

For banks, capital inflow is an interest income, and the income function is as follows:

$$\max_{\tau^{BC}} \pi_b^{BC} = \left[\left(c_s^{BC} + b_s^{BC}\right) D^{BC} + k^{BC} e^2/2 + \gamma^{BC} a^2/2\right] \tau^{BC} \tag{21}$$

Here, $D^{BC} = (1-\beta)q - p^c + a\xi + \mu e + \epsilon$.

Suppliers and e-commerce platforms are both risks averse. By using the mean - standard deviation to characterize the risk-averse behavior of enterprise managers, we can get the utility functions of suppliers and e-commerce platforms as follows: $U\left(\pi_e^{BD}\right) = \left(x^{BD} p_s^{BD} - c_e^{BD} - b_e^{BD}\right)\left[\beta q - p^d + a\xi + \mu e - \eta\sigma\right] - \gamma^{BD} a^2/2$. It is easy to prove that $U\left(\pi_s^{BD}\right)$ and $U\left(\pi_e^{BC}\right) = \left(p_e^{BC} - w_d^{BC} - b_e^{BC}\right)\left[(1-\beta)q - p^c + a\xi + \mu e - \eta\sigma\right] - \gamma^{BC} a^2/2$. By solving the first-order derivative and the second-order derivative, it can be concluded that $U\left(\pi_s^{BC}\right)$ and $U\left(\pi_e^{BC}\right)$ are concave functions, so there is an optimal solution.

Proposition 6: Under blockchain technology, according to Equations 18–21, it can be calculated that, the optimal retail price of the green e-commerce platform to green preference consumers satisfies $p^{BC^*} = \frac{b_e + w_d + a \quad \xi - \sigma n + e \quad \mu + q \quad (1-\beta)}{2}$, and the optimal price of the manufacturer is $w_d^{BC^*} = \frac{b_d + c_d + w_s + a \quad \xi - \sigma n + e \quad \mu - q \quad (\beta - 1)}{2}$. The optimal price of the supplier satisfies $w_s^{BC^*} = \frac{(\tau+1) \quad (2 \quad ke + b \quad \mu + c_s \quad \mu)}{\mu}$, the optimal trust is $a^{BC^*} = \frac{\xi(b + w_d + \sigma\eta - e\mu + q \quad (\beta - 1))}{\xi^2 - 2\gamma}$, and the optimal green preference is $e^{BC^*} = \frac{b - q + w_d - a \quad \xi + \eta\sigma + q \quad \beta}{\mu}$. The optimal expected financing utility of the supplier is $\pi_s^{BC^*} = (\tau + 1) \quad \left(\frac{k \quad (b - q + w_d - a\xi + \eta\sigma + q\beta)^2}{2 \quad \mu^2} + 30 \quad a^2\right)$, the optimal expected financing utility of the manufacturer is $\pi_d^{BC^*} = \frac{w_d(2b + 2c_d - w_d + 2w_s)}{8} - 30 \quad a^2$, and the optimal expected financing utility of the green e-commerce is $\pi_e^{BC^*} = \frac{(w_d - w_s - b - c_d)(b + c_d + w_s)}{4} + 30 \quad a^2$.

Proof: The optimal solution of each variable is solved by backward induction of Stackelberg game, and the derivation process is like Theorem 2.

Inference 6 Under the blockchain, technology can reduce the credit risk as much as possible. Smart contracts and consensus mechanisms can effectively ensure that suppliers repay loans on time and meet the risk aversion of financial institutions. The optimal retail price p_e^{BC} is positively correlated with the cost of blockchain b_e^{BC}, will increase with the increase of the Consignment sales market size $(1-\beta)q$, and is negatively correlated with the risk aversion η^{BC}. The manufacturer's optimal price w_d^{BC} is positively correlated with the Consignment sales scale $(1-\beta)q$ and will increase with the increase of product greenness e^{BC}. The supplier's price w_s^{BC} is positively correlated with the cost of blockchain b_s^{BC} and will rise with the increase of interest rates τ. The optimal trust a is positively correlated with blockchain technology b and risk aversion coefficient η, but will decrease with the increase of product greenness e. The optimal product greenness e will also increase with the increase of blockchain technology b, attracting more consumers to recover funds.

4.4 Comparative analysis

Blockchain technology facilitates collaboration between different actors, allowing all parties in the e-commerce supply chain to exchange information in a shared, trusted network. This helps break down information silos, form blockchain trust, and improve overall operational efficiency. With blockchain technology, a faster and more transparent financing process can be achieved. Using blockchain smart contracts to automate many steps in the financing process not only reduces wait times but also reduces the risk of human intervention. The use of blockchain technology can reduce intermediaries and the reliance on traditional paper documents, reduce financing costs while improving the speed of financing, which not only reduces processing time and costs, but also improves the accuracy of data. In general, after the introduction of blockchain, the income level of suppliers, manufacturers and e-commerce enterprises has been significantly improved. Although the introduction of blockchain will increase the cost of blockchain adoption compared to the situation without the introduction of blockchain technology, it will also save part of the cost of information review and improve the financing efficiency of e-commerce supply

chain banks. The introduction of blockchain technology has further deepened the degree of information sharing and promoted the transfer of trust. The introduction of blockchain technology helps enterprises better avoid risks.

By analyzing the previous model, we can further draw conclusions.

Proposition 7: Under Direct Selling, when the supplier's production cost is met $c_s < b_s(1+\tau)/(\delta+\varepsilon)$, the greenness of the product under the blockchain trust mechanism is higher than that without blockchain.

Proof: According to the optimal solution of decision variables in theorem 3 and theorem 5, it can be seen that under Direct Selling: $e^{PD^*} = \frac{\mu\ (c_s+c_s\ \ \delta+c_s\ \ \varepsilon-a\xi+c_s\ \ \tau+\eta\sigma-\beta q+a\ \ \xi x-\eta\sigma\ \ x+x\ \ \beta q)}{2\ \ k+2\ \ \delta\ \ k+2\ \ k\varepsilon+2\ \ k\tau+\mu^2\ \ x-\mu^2}$, $e^{BD^*} = \frac{\mu\ \ (b_s+b_s\ \ \tau+c_s-a\ \ \xi+c_s\tau+\sigma\eta-q\beta+a\ \ \xi x-\sigma\ \ \eta x+q\ \ x\beta)}{2\ \ k+2\ \ k\tau+\mu^2\ \ x-\mu^2}$.When $c_s < b_s(1+\tau)/(\delta+\varepsilon)$,that is, $c_s\ \ (\delta+\varepsilon) < b_s(1+\tau)$, we get $\mu\ \ (b_s+b_s\ \ \tau+c_s-a\ \ \xi+c_s\tau+\sigma\eta-q\beta+a\ \ \xi x-\sigma\ \ \eta x+q\ \ x\beta) > \mu\ \ (c_s+c_s\ \ \delta+c_s\ \ \varepsilon-a\xi+c_s\ \ \tau+\eta\sigma-\beta q+a\ \ \xi x-\eta\sigma\ \ x+x\ \ \beta q)$, and $2k+2k\tau+\mu^2x-\mu^2 < 2k+2\delta k+2k\varepsilon+2k\tau+\mu^2x-\mu^2$, therefore $e^{BD^*} > e^{RD^*}$.

The adoption of blockchain has fundamentally alleviated the financing constraints of enterprises. Under blockchain technology, e-commerce enterprises are more willing to provide guarantees for cash-strapped suppliers, financial institutions are more willing to provide chain financial services, and loan services have greatly solved the funding problem. In the process of supply chain operation, when the cost meets certain conditions under Direct Selling, compared with the non-blockchain technology, the introduction of blockchain technology improves the greenness of the product, helps attract more consumers who prefer green, and quickly alleviates their own financial constraints.

Proposition 8: In Consignment consignment, when the supplier's production cost meets $c_s < \frac{k\ \ e^2}{b_e+w_d+a\ \ \xi}$, under the block chain trust mechanism, the supplier will reduce the selling price to attract more consumers and obtain more profits.

Proof: According to theorem 4 and theorem 6, we know $p_e^{RC^*} = \frac{k\ \ e^2+2\ \ c_s\mu\ \ e+2\ \ c_s\ \ q-2\ \ c_s\ \ \eta\sigma-2\ \ c_s\ \ q\beta}{2\ \ c_s} = \frac{k\ \ e^2+2\ \ c_s[\mu\ \ e-\eta\sigma+(1-\beta)q]}{2\ \ c_s}$, $p^{BC^*} = \frac{b_e+w_d+a\ \ \xi-\sigma n+e\ \ \mu+q\ \ (1-\beta)}{2}$. When $c_s < \frac{k\ \ e^2}{b_e+w_d+a\ \ \xi}$, $b_e+w_d+a\ \ \xi < \frac{k\ \ e^2}{c_s}$, therefore $p_e^{RC^*} > p_e^{BC^*}$.

The introduction of blockchain technology is crucial to the financing decision of the e-commerce supply chain, which can better allow financial constraint suppliers to make scientific and reasonable decisions. Blockchain increases trust in the e-commerce supply chain sales process and enhances the traceability of information. In Consignment, blockchain technology has improved the transparency of information. In order to attract consumers, it will reduce the selling price, improve the competitiveness of enterprises and ease the financial constraint.

5 Numerical analysis

To further validate the model, MATLAB is used to analyze the expressions and explore the impact of blockchain, blockchain trust mechanisms, risk avoidance, product greenness, guarantee rates, and information verification efforts on bank finance. Based on Wu et al. [38] and An et al. [56], the parameters for each decision variable are follows: $a = 0.6, \beta = 0.5, q = 100, c_s = 0.6, c_d = 0.8, \tau = 0.06, x = 0.05, \xi = 10$. The numerical settings meet the proof process of the equations to ensure that the decision-making process.

5.1 Influence of blockchain technology

The impact of blockchain on green e-commerce supply chain bank financing strategies is shown in Figure 2. Figure 2a illustrates the change in supplier profits with the blockchain under the Direct Selling and Consignment sales models. Figure 2b shows the impact on green e-commerce.

For suppliers, the financing benefits under both the Direct Selling and Consignment sales models decrease as the blockchain adoption cost increases, with a larger reduction under the Consignment model. The impact on the Direct Selling model is relatively small. In the Direct Selling model, supplier utility decreases with the increase in blockchain adoption, but the change is not significant, indicating that blockchain has a small impact. In the Consignment model, the introduction of blockchain causes a significant drop in supplier benefits because it increases the financing costs, which negatively affects the financing returns. However, when the cost of blockchain technology $b < 1.45$, suppliers should be more inclined to use the Consignment method, because the Consignment method can create better benefits for suppliers and ease financial constraints. When the cost of blockchain technology $b > 1.45$, Direct Selling should be adopted.

For e-commerce platforms, the impact of blockchain on platforms varies greatly between the Direct Selling and Consignment models. In the Direct Selling model, the introduction of blockchain technology can attract more consumers, increase sales, and improve the platform's overall profitability. The platform charges a percentage of service fees, and the increased sales offset the increased blockchain adoption cost, has a positive effect on profits. In the Consignment model, however, the platform's profits decrease as the blockchain technology cost rises, and the decline is more pronounced. When the blockchain technology cost $b < 6.79$, the Consignment model can generate more profit for the platform, helping to better diversify its risks. When blockchain technology cost $b > 6.79$, Direct Selling should be chosen.

Overall, when blockchain costs are low, both suppliers and green e-commerce platforms are likely to choose the Consignment sales model, as it better alleviates the supplier's financial constraints and provides better returns for both parties. When the blockchain technology cost is high, the Direct Selling model becomes the better option.

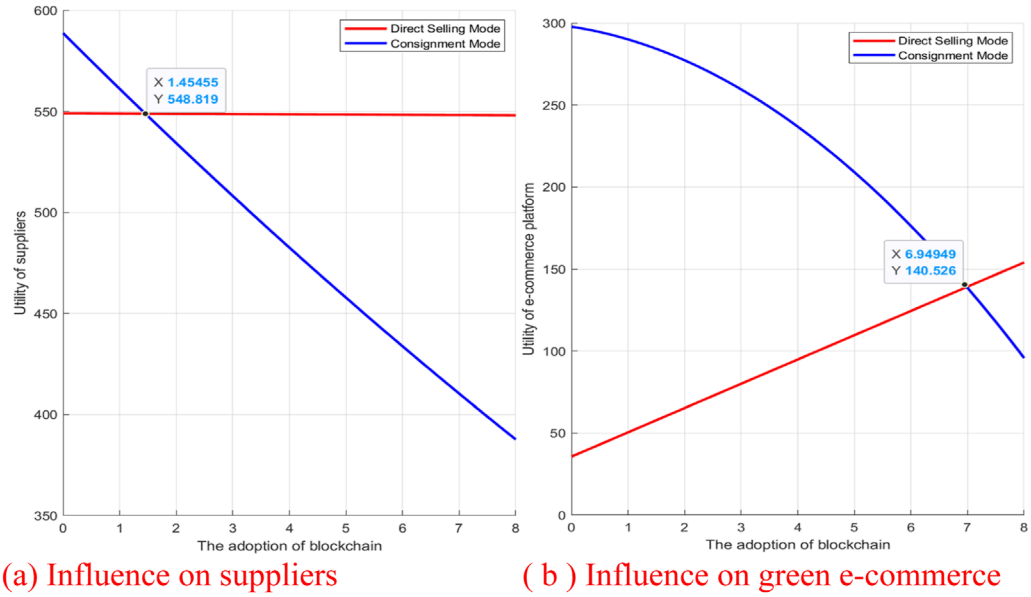


FIGURE 2
Influence of blockchain. **(a)** Influence on suppliers. **(b)** Influence on green e-commerce.

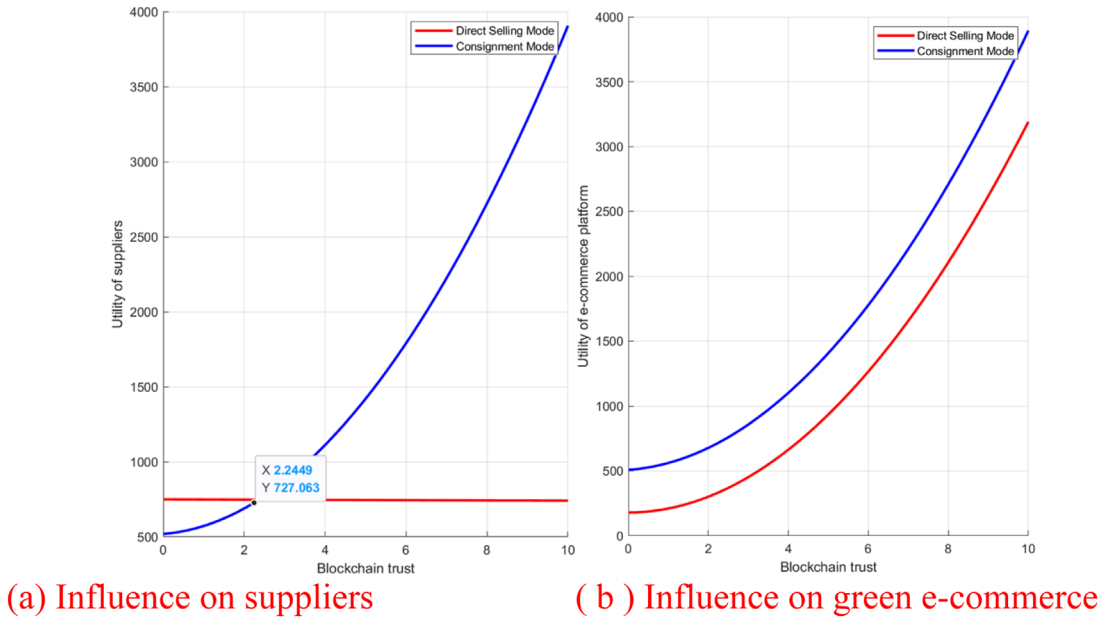


FIGURE 3
Influence of blockchain trust. **(a)** Influence on suppliers. **(b)** Influence on green e-commerce.

5.2 Influence of blockchain trust

The impact of blockchain trust on bank financing strategies is shown in Figure 3. Figure 3a illustrates the trend of supplier benefits as blockchain trust increases, while Figure 3b shows the impact on the green e-commerce platform.

For suppliers, blockchain trust has a very different impact on supplier efficiency under Direct Selling and Consignment. Under Direct Selling, the vendor's revenue will decrease slightly as the blockchain trust increases. Blockchain trust has less impact on suppliers and has a limited effect on easing financial constraints. In Consignment, blockchain trust has a great impact on the profit of suppliers, so it can better obtain more profits and obtain a good profit level, which helps to alleviate the fund problem of the supply chain. With the increase of

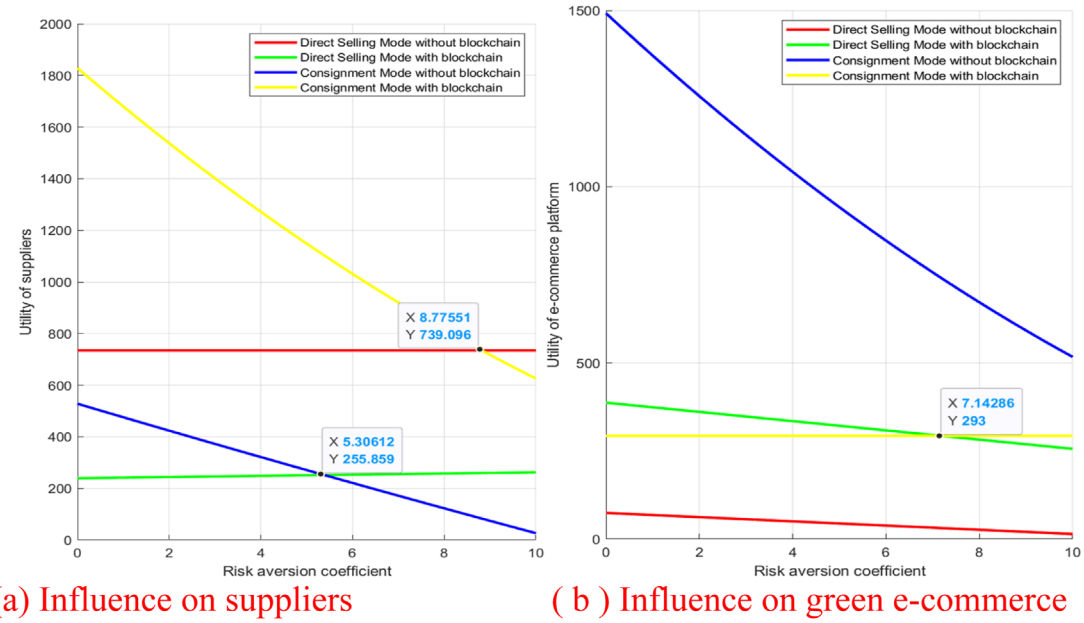


FIGURE 4
Influence of risk aversion. **(a)** Influence on suppliers. **(b)** Influence on green e-commerce.

blockchain trust, the revenue of suppliers shows a good growth trend, and the rise is large, and there is a good room for development. When the blockchain trust level a<2.24, Direct Selling can bring better returns to suppliers. However, when the blockchain trust exceeds a certain threshold, suppliers will be more willing to introduce the blockchain trust in the Consignment consignment to improve their overall income. In practice, with a high degree of blockchain trust, enterprises should actively adopt the Consignment selling method to better play the role of blockchain trust mechanism in financing strategy optimization.

For e-commerce platforms, the impact of blockchain trust on platform profits shows little difference between the Direct Selling and Consignment models. In the Direct Selling model, platform profits increase significantly with the increase in blockchain trust, with considerable fluctuation. The higher the blockchain trust, the higher the platform's profitability. In the Consignment model, as blockchain trust increases, the platform's profits show a clear upward trend, and blockchain trust helps enhance platform profits and fosters better development. In both sales models, the Consignment model generates better benefits for e-commerce platforms compared to the Direct Selling model.

At a certain level of blockchain trust, both suppliers and e-commerce platforms tend to choose the Consignment model. In this model, when blockchain trust is low, the effect of blockchain trust on supplier and platform benefits is minimal, and the difference between the two sales models is small. However, as blockchain trust increases, the impact grows, and the benefits brought to e-commerce supply chain enterprises become more significant.

5.3 Influence of risk aversion

The impact of the risk aversion on suppliers and green e-commerce platforms is shown in Figure 4. This analysis compares the effects of risk aversion on bank finance strategies before and after adopting blockchain in the Direct Selling and Consignment sales models. Figure 4a shows the impact on suppliers, and Figure 4b shows the impact on green e-commerce platforms.

For suppliers, under Direct Selling and before the adoption of blockchain technology, the risk avoidance coefficient has little impact on suppliers. After the adoption of blockchain technology, the financing returns of suppliers will increase slightly with the increase of risk aversion coefficient. The profit before the adoption of blockchain is higher than that after the adoption of blockchain, indicating that blockchain has a greater effect on the relationship between risk avoidance and supplier benefits. In Consignment consignment, before the use of blockchain, the benefit of suppliers will gradually decrease with the increase of risk avoidance coefficient. After the adoption of blockchain, the risk avoidance coefficient will also cause the supplier's income to decline, and the decline is large. In the case of blockchain, the supplier benefit decreases more with the risk avoidance coefficient than without blockchain, and blockchain increases the impact of risk avoidance behavior on supplier financing. Suppliers should control their risk avoidance behavior within a reasonable range. At a low risk aversion factor, Direct Selling should be actively adopted when not using blockchain technology. When using blockchain technology, suppliers should use

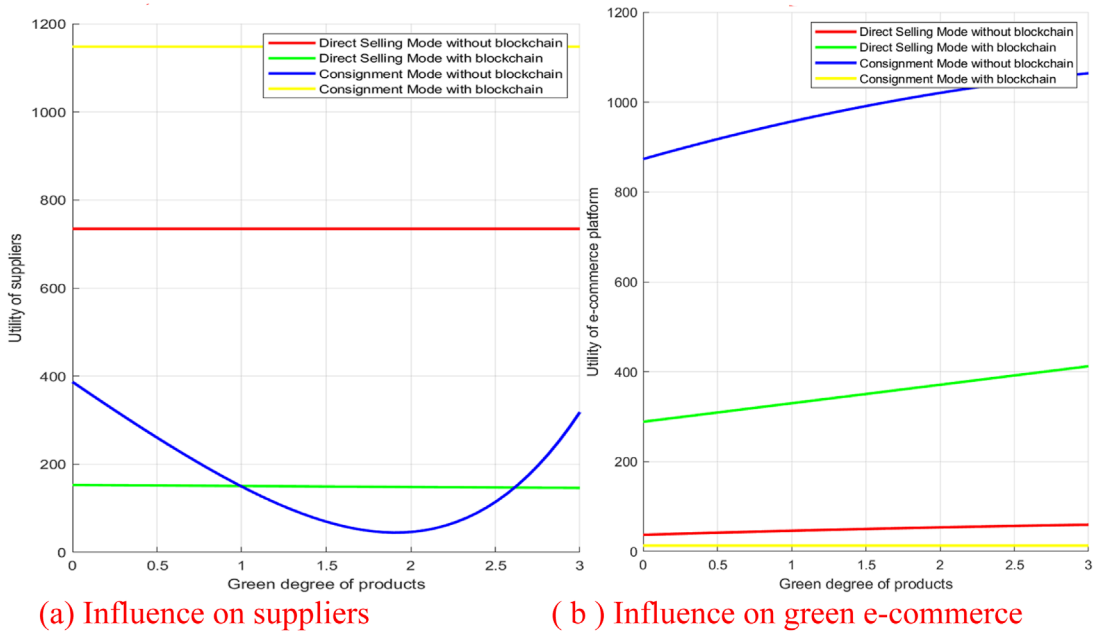


FIGURE 5
Influence of product greenness. **(a)** Influence on suppliers. **(b)** Influence on green e-commerce.

Consignment. When suppliers have high risk aversion, they should give priority to Direct Selling, and use blockchain technology to enhance credit transparency and reduce additional financing costs caused by information asymmetry.

For the e-commerce platform, in the Direct Selling mode, before the adoption of blockchain technology, the utility of the e-commerce platform will decrease with the increase of risk avoidance coefficient, and a high-risk avoidance coefficient is not conducive to the development of the e-commerce platform. After the adoption of blockchain, the high-risk avoidance coefficient will also lead to the reduction of the utility of the e-commerce platform. Before and after the adoption of blockchain, the utility of e-commerce platforms has a small variation with the risk avoidance coefficient. In Consignment consignment, before the use of blockchain, the risk avoidance coefficient will cause the utility of e-commerce platform to decline sharply, and the change range is larger than that of Direct Selling. After the adoption of blockchain, the risk avoidance coefficient has little impact on the financing utility of e-commerce platforms. Compared to the non-blockchain scenario, the adoption of blockchain reduces the impact of risk avoidance coefficient on the utility of e-commerce platforms. Under the lower risk avoidance coefficient, when the block chain technology is not used, the level desk is more inclined to use Consignment, which can obtain higher financing returns. With the adoption of blockchain technology, Direct Selling will be a better option. Under the high-risk avoidance coefficient, before and after the use of blockchain technology, Consignment can bring better returns to the e-commerce platform.

In the green e-commerce supply chain, a higher risk aversion coefficient does not necessarily contribute to increasing financing utility for suppliers and e-commerce platforms. Managers should appropriately reduce their risk aversion tendencies in order to better assist suppliers in obtaining financing.

5.4 Influence of product greenery

The impact of product greenery on bank financing strategies is shown in Figure 5. Figure 5a shows the impact on suppliers, and Figure 5b shows the impact on e-commerce platforms.

For suppliers, in the Direct Selling model, prior to the blockchain, product greenery has a minimal effect on supplier financing utility. After adopting blockchain technology, a certain degree of product greenery leads to a slight decrease in supplier utility, as higher product greenery increases the supplier's green investment costs. In the Consignment model, prior to blockchain adoption, supplier financing utility decreases initially with the increase in product greenery but later increases. After adopting blockchain, the effect of product greenery on supplier financing utility becomes less significant. The introduction of a blockchain increases supplier financing benefits in the Consignment model but reduces them in the Direct Selling model. Companies should actively improve the greenness of their products and increase the adoption of blockchain technology.

For e-commerce platforms, in the Direct Selling model, prior to adopting blockchain technology, platform profits increase as product greenery increases, as higher product greenery attracts more consumers and increases platform profits. After blockchain adoption, product greenery continues to boost platform profits. Compared to the non-blockchain scenario, the blockchain increases the influence of product

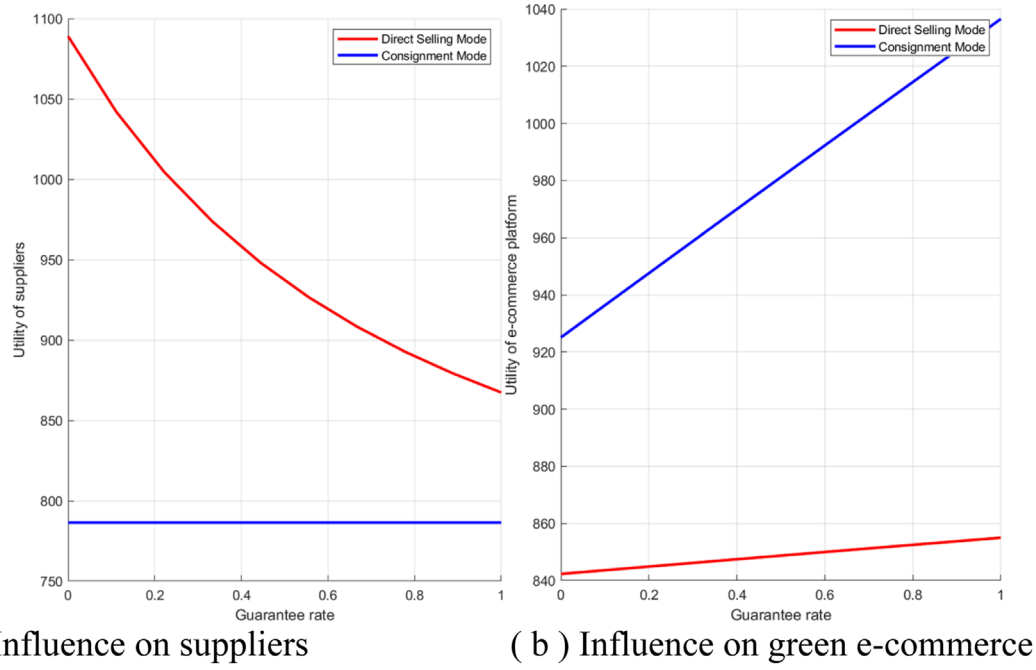


FIGURE 6
Influence of guaranteed rate. **(a)** Influence on suppliers. **(b)** Influence on green e-commerce.

greenery on platform profits. In the Consignment model, prior to blockchain adoption, platform financing profits increase significantly as product greenery increases. After adopting blockchain, the impact of product greenery on platform profits becomes limited. The adoption of blockchain reduces the influence of product greenery on platform profits.

A certain level of product greenery attracts more green-preference consumers, helping to increase platform profits, but it increases costs for suppliers. Only after exceeding a certain threshold does the non-blockchain-based Consignment model boost supplier financing benefits.

5.5 Influence of guarantee rate

The impact of the guaranteed rate on suppliers and green e-commerce is shown in Figure 6. Figure 6a is the influence on suppliers, and Figure 6b is the influence on green e-commerce.

For suppliers, in the Direct Selling model, supplier financing benefits decrease as the guaranteed rate increases. A higher guaranteed rate raises the supplier's costs, and while it helps solve funding issues, it also increases the financing cost burden. However, in the Consignment model, the guaranteed rate has a minimal influence on the overall revenue and does not reduce supplier utility while helping alleviate funding constraints. Although the Direct Selling model leads to a decrease in supplier financing benefits, overall, the revenue level remains higher than that in the Consignment model.

For green e-commerce, an increase in the guaranteed rate promotes an increase in platform profits in both the Direct Selling and Consignment models. In the Consignment model, the effect of the guaranteed rate on platform profits is significant, with a large growth margin. In the Direct Selling model, platform utility also increases with an increase in the guaranteed rate, but the growth is slower, and the increase is smaller. E-commerce is likely the Consignment model, which yields higher profits compared to the Direct Selling model.

The impact of the guaranteed rate on suppliers and e-commerce platforms is quite different. An increase in the guaranteed rate increases supplier financing costs but promotes the increase of e-commerce platform profits. The guaranteed rate also influences the choice of sales model. Suppliers are more likely to choose the Direct Selling model, while e-commerce platforms prefer the Consignment model.

5.6 Influence of information verification effort

The impact of corporate information verification efforts on e-commerce supply chain bank financing strategies is shown in Figure 7. Figure 7a represents the impact on suppliers, and Figure 7b represents the impact on green e-commerce.

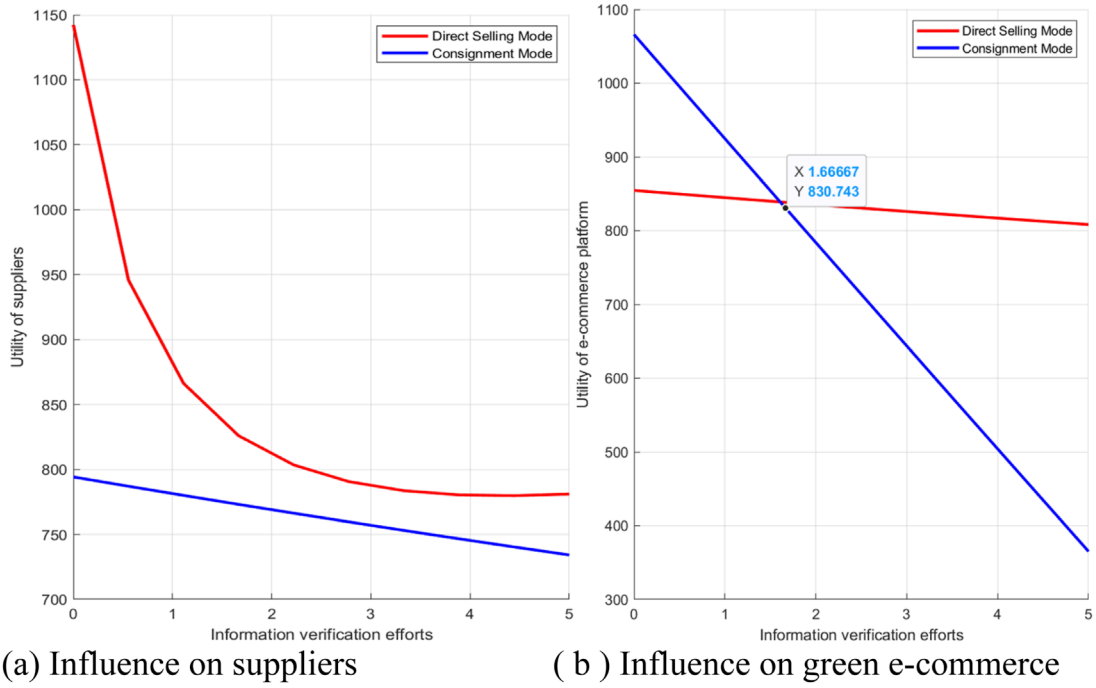


FIGURE 7
Influence of information verification efforts. **(a)** Influence on suppliers. **(b)** Influence on green e-commerce.

For suppliers, information verification efforts increase costs, and in both the Direct Selling and Consignment models, supplier benefits decrease as information verification efforts increase. Supplier benefits are higher in Direct Selling than in Consignment. When information verification efforts are within a certain threshold, the decrease in supplier benefits is larger in Direct Selling compared to the Consignment.

For e-commerce platforms, utility decreases as information verification efforts increase. In the Consignment model, the impact of information verification efforts on platform utility is larger, and the decrease is more noticeable compared to the Direct Selling model. At lower levels of information verification efforts, the Consignment model creates more profit for the e-commerce platform.

Information verification efforts lead to a decrease in financing benefits for both suppliers and e-commerce platforms. For suppliers, the decrease is more significant in the Direct Selling model, while for e-commerce platforms, the decrease is larger in the Consignment model. Considering the impact of information verification efforts, the choice of sales model will differ for the two parties.

6 Conclusion

Blockchain opens up new pathways for optimizing e-commerce supply chain bank financing strategies, overcoming the financing difficulties faced by enterprises in practice, and effectively solving the financing challenges encountered by suppliers, improving financing efficiency. To solve the challenges of difficult and expensive financing, this study explores the influence of blockchain, blockchain trust, risk aversion, product greenery, guarantee rates, and information verification efforts on bank financing strategies in e-commerce supply chains, combining the Direct Selling and Consignment sales models, and considering enterprise risk aversion behavior and consumer green preferences. The research shows that when blockchain technology is within a certain threshold, suppliers are more likely to choose the Consignment model, which helps alleviate their funding constraints. Blockchain trust mechanisms have a limited impact on financing strategies in the Direct Selling model, but they can provide better financing benefits to suppliers in the Consignment model. A higher risk aversion coefficient does not necessarily contribute to improving finance efficiency, and its impact on suppliers is greater with blockchain adoption than without, a higher risk aversion coefficient will improve the trust of blockchain. A certain degree of product generator increases supplier costs but helps attract more consumers and boosts overall revenue. Information verification efforts and higher guarantee rates increase supplier financing costs, with different impacts on the Direct Selling and Consignment sales models.

In the e-commerce supply chain, our work provides some targeted management implications and novel insights for the financially constrained e-commerce supply chain. In the process of bank financing, blockchain technology can effectively improve information transparency, and the e-commerce supply chain should actively introduce blockchain technology. When the adoption cost of blockchain technology is low, the proportion of Consignment sales products should be increased as much as possible, which helps to withdraw funds quickly to alleviate capital constraints. Suppliers can increase the market scale of Consignment method, to better exert the influence of

blockchain trust mechanism on financing strategy, increase the trust between enterprises and between enterprises and consumers, and help suppliers obtain more profits to improve financing efficiency. E-commerce supply chain enterprises can appropriately reduce their risk avoidance tendency, and higher risk avoidance is not conducive to the increase of financing returns. Blockchain technology will expand the impact of risk avoidance coefficient, when the enterprise has a high-risk aversion tendency, should be careful to choose to adopt blockchain technology. High risk aversion means that enterprises are more conservative in the financing process, which may lead to less efficient use of funds, so when adopting new technologies or adjusting product strategies, they need to make a trade-off between reducing risks and obtaining higher returns. Suppliers should actively produce green and low-carbon products. Increasing green investment in products may increase production costs, but at the same time, it can obtain higher market premium and consumer trust, thus improving financing efficiency. Considering the impact of guaranteed interest rate and information verification efforts on suppliers' financing returns, Direct Selling is more helpful to ease suppliers' financial constraints.

Although this study provides in-depth research on blockchain-based financing, it is important to acknowledge its limitations. Specifically, the influence of blockchain may involve multiple layers of mechanisms, and the application of different blockchain characteristics in financing deserves further exploration. Additionally, our study only utilizes simple random market demand, which may not fully capture the real market environments. Future research can explore more market-specific random demand that reflects market uncertainties. Future studies may also examine the influence of enterprise behavior on e-commerce supply chain financing efficiency, combining more comprehensive enterprise decision-making behavior to expand related research. This study only considers a single bank financing model, and platform financing of the future e-commerce supply chain can be used as an alternative financing scheme. The selection of multiple financing methods is also a topic worthy of study.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

LY: Methodology, Writing–original draft, Writing–review and editing. QH: Conceptualization, Project administration, Supervision, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Liaoning Provincial Department of Education's basic scientific research project in 2023, "Research on the path of matching supply and demand of smart pension industry under government subsidy mechanism" (JYTMS20231228).

Acknowledgments

We acknowledge all those who contributed to this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Zhang L, Cui L, Chen L, Dai J, Jin Z, Wu H. A hybrid approach to explore the critical criteria of online supply chain finance to improve supply chain performance. *Int J Prod Econ* (2023) 255:108689. doi:10.1016/j.ijpe.2022.108689

2. Wuttke DA, Rosenzweig ED, Heese HS. An empirical analysis of supply chain finance adoption. *J Operations Manage* (2019) 65(3):242–61. doi:10.1002/joom.1023

3. Jia F, Zhang T, Chen L. Sustainable supply chain Finance: towards a research agenda. *J Clean Prod* (2020) 243:118680. doi:10.1016/j.jclepro.2019.118680

4. Mandal P, Basu P, Choi T-M, Rath SB. Platform financing vs. bank financing: strategic choice of financing mode under seller competition. *Eur J Oper Res* (2024) 315(1):130–46. doi:10.1016/j.ejor.2023.11.025

5. Chang S, Li A, Wang X, Wang X. Joint optimization of e-commerce supply chain financing strategy and channel contract. *Eur J Oper Res* (2022) 303(2):908–27. doi:10.1016/j.ejor.2022.03.013

6. Yu H, Zhao Y, Liu Z, Liu W, Zhang S, Wang F, et al. Research on the financing income of supply chains based on an E-commerce platform. *Technol Forecast Soc Change* (2021) 169:120820. doi:10.1016/j.techfore.2021.120820

7. Chen Z, Chen J, Zhang Z, Zhi X. Does network governance based on banks'e-commerce platform facilitate supply chain financing? *China Agric Econ Rev* (2019) 11(4):688–703. doi:10.1108/caer-06-2018-0132

8. Zhou L, Wang J, Li F, Xu Y, Zhao J, Su J. Risk aversion of B2C cross-border e-commerce supply chain. *Sustainability* (2022) 14, 8088. doi:10.3390/su14138088

9. Lan C, Zhu J. New product presale strategies considering consumers' loss aversion in the E-commerce supply chain. *Discrete Dyn Nat Soc* (2021) 2021(1):1–13. doi:10.1155/2021/8194879

10. Du J, Shi Y, Li W, Chen Y. Can blockchain technology be effectively integrated into the real economy? Evidence from corporate investment efficiency. *China J Account Res* (2023) 16(2):100292. doi:10.1016/j.cjar.2023.100292

11. Yang H, Zhen Z, Yan Q, Wan H. Mixed financing scheme in a capital-constrained supply chain: bank credit and e-commerce platform financing. *Int Trans Oper Res* (2022) 29(4):2423–47. doi:10.1111/itor.13090

12. Shi P, Wang J, Hu Y, Yin H, Chen Z, Xu B, et al. Heterogeneous products operation decisions of online dual-channel supply chain considering online reviews under different financing modes. *RAIRO-Operations Res* (2024) 58(5):4013–33. doi:10.1051/ro/2024144

13. Zhao S, Lu X. Accounts receivable financing and supply chain coordination under the government subsidy. In: *Paper presented at the 2021 11th International Conference on information science and technology (ICIST)* (2021).

14. Hagendorff J, Lim S, Nguyen DD. Lender trust and bank loan contracts. *Manage Sci* (2023) 69(3):1758–79. doi:10.1287/mnsc.2022.4371

15. Peng Q, Wang C, Goh M. Green financing strategies in a low-carbon e-commerce supply chain under service quality regulation. *Environ Sci Pollut Res* (2023) 30(2):2575–96. doi:10.1007/s11356-022-22329-w

16. Tang Q, Lu Y, Wang B, Li Z. A deep Convolutional neural network based risk identification method for E-commerce supply chain finance. *Scientific Programming* (2022) 2022(1):1–10. doi:10.1155/2022/6298248

17. Bai S, Jia X. Agricultural supply chain financing strategies under the impact of risk attitudes. *Sustainability* (2022) 14(14):8787. doi:10.3390/su14148787

18. Yan N, Sun B, Zhang H, Liu C. A partial credit guarantee contract in a capital-constrained supply chain: financing equilibrium and coordinating strategy. *Int J Prod Econ* (2016) 173:122–33. doi:10.1016/j.ijpe.2015.12.005

19. Yan B, Chen Z, Yan C, Zhang Z, Kang H. Evolutionary multiplayer game analysis of accounts receivable financing based on supply chain financing. *Int J Prod Res* (2024) 62(22):8110–28. doi:10.1080/00207543.2021.1976432

20. Lou Z, Xie Q, Shen JH, Lee C-C. Does supply chain finance (SCF) alleviate funding constraints of SMEs? Evidence from China. *Res Int Business Finance* (2024) 67:102157. doi:10.1016/j.ribaf.2023.102157

21. Ma H-L, Wang Z, Chan FT. How important are supply chain collaborative factors in supply chain finance? A view of financial service providers in China. *Int J Prod Econ* (2020) 219:341–6. doi:10.1016/j.ijpe.2019.07.002

22. Bordino I, Gullo F, Legnaro G. Advancing receivable financing via a network-based approach. *IEEE Trans Netw Sci Eng* (2020) 8(2):1328–37. doi:10.1109/tnse.2020.3005612

23. Yao X, Li X, Mangla SK, Song M. Roles of AI: financing selection for regretful SMEs in e-commerce supply chains. *Transportation Res E: Logistics Transportation Rev* (2024) 189:103649. doi:10.1016/j.tre.2024.103649

24. Liu L, Peng Q. Evolutionary game analysis of enterprise green innovation and green financing in platform supply chain. *Sustainability* (2022) 14(13):7807. doi:10.3390/su14137807

25. Cao B-B, You T-H, Ou CX, Zhu H, Liu C-Y. Optimizing payment schemes in a decentralized supply chain: a Stackelberg game with quality investment and bank credit. *Comput and Ind Eng* (2022) 168:108077. doi:10.1016/j.cie.2022.108077

26. Zhou W, Lin T, Cai G. Guarantor financing in a four-party supply chain game with leadership influence. *Prod Operations Manage* (2020) 29(9):2035–56. doi:10.1111/poms.13196

27. Wang L, Zhang L, Zhang X. Optimal financing strategies for low-carbon supply chains: a Stackelberg game perspective. *Managerial Decis Econ* (2025) 46(1):666–80. doi:10.1002/mde.4394

28. Zhang Z-c, Xu H-y, Chen K-b. Operational decisions and financing strategies in a capital-constrained closed-loop supply chain. *Int J Prod Res* (2021) 59(15):4690–710. doi:10.1080/00207543.2020.1770356

29. Dong G, Liang L, Wei L, Xie J, Yang G. Optimization model of trade credit and asset-based securitization financing in carbon emission reduction supply chain. *Ann Operations Res* (2023) 331:35–84. doi:10.1007/s10479-021-04011-5

30. Abdel-Basset M, Mohamed R, Sallam K, Elhoseny M. A novel decision-making model for sustainable supply chain finance under uncertainty environment. *J Clean Prod* (2020) 269:122324. doi:10.1016/j.jclepro.2020.122324

31. Guo J, Jia F, Yan F, Chen L. E-commerce supply chain finance for SMEs: the role of green innovation. *Int J Logistics Res Appl* (2024) 27(9):1596–615. doi:10.1080/13675567.2023.2167959

32. Ma S, Qian Q, Wang G, Xu M. Research on smart contracts of accounts receivable financing in supply chain finance based on blockchain technology. In: *Paper presented at the 2023 4th International Conference on E-commerce and Internet technology (ECIT 2023)* (2023).

33. Li J, Zhu S, Zhang W, Yu L. Blockchain-driven supply chain finance solution for small and medium enterprises. *Front Eng Manage* (2020) 7(4):500–11. doi:10.1007/s42524-020-0124-2

34. Wang D, Zhao D, Chen F. Research on financing strategy of green energy-efficient supply chain based on blockchain technology. *Energies* (2023) 16(7):2985. doi:10.3390/en16072985

35. Rijanto A. Blockchain technology adoption in supply chain finance. *J Theor Appl Electron Commerce Res* (2021) 16(7):3078–98. doi:10.3390/jtaer16070168

36. Su L, Cao Y, Li H, Tan J. Blockchain-driven optimal strategies for supply chain finance based on a tripartite game model. *J Theor Appl Electron Commerce Res* (2022) 17(4):1320–35. doi:10.3390/jtaer17040067

37. Tang D, Zhuang X. Financing a capital-constrained supply chain: factoring accounts receivable vs a BCT-SCF receivable chain. *Kybernetes* (2020) 50(8):2209–31. doi:10.1108/k-06-2020-0367

38. Wu C, Xu C, Zhao Q, Zhu J. Research on financing strategy under the integration of green supply chain and blockchain technology. *Comput and Ind Eng* (2023) 184:109598. doi:10.1016/j.cie.2023.109598

39. Li Q, Ma M, Shi T, Zhu C. Green investment in a sustainable supply chain: the role of blockchain and fairness. *Transportation Res Part E: Logistics Transportation Rev* (2022) 167:102908. doi:10.1016/j.tre.2022.102908

40. Xia Y, Shang R, Wei M, Wei Z. Optimal pricing and financing decision of dual-channel green supply chain considering product differentiation and blockchain. *Ann Operations Res* (2024) 338(2):1317–58. doi:10.1007/s10479-024-05996-5

41. Hao M, Ding J. Decision-making and impact of blockchain on accounts receivable financing. In: *Paper presented at the LISS 2021: Proceedings of the 11th International Conference on logistics, Informatics and service Sciences* (2022).

42. Chen J, Chen S, Liu Q, Shen M. Applying blockchain technology to reshape the service models of supply chain finance for SMEs in China. *Singapore Econ Rev* (2021) 1–18. doi:10.1142/s0217590821480015

43. Jiang R, Kang Y, Liu Y, Liang Z, Duan Y, Sun Y, et al. A trust transitivity model of small and medium-sized manufacturing enterprises under blockchain-based supply chain finance. *Int J Prod Econ* (2022) 247:108469. doi:10.1016/j.ijpe.2022.108469

44. Shu C, Chen Y, Tan C, Luo Y, Dou H. Enhancing trust transfer in supply chain finance: a blockchain-based transitive trust model. *J Cloud Comput* (2024) 13(1):4. doi:10.1186/s13677-023-00557-w

45. Jiang Y, Li Y, Wei J, Liu Y. Research on trust mechanism of supply chain finance under Industrial Internet embedded with blockchain. *Plos one* (2024) 19(6):e0299011. doi:10.1371/journal.pone.0299011

46. Tsai C-H. Supply chain financing scheme based on blockchain technology from a business application perspective. *Ann Operations Res* (2023) 320(1):441–72. doi:10.1007/s10479-022-05033-3

47. Sun R, He D, Su H. Evolutionary game analysis of blockchain technology preventing supply chain financial risks. *J Theor Appl Electron Commerce Res* (2021) 16(7):2824–42. doi:10.3390/jtaer16070155

48. Nozari H, Ghahremani-Nahr J. Assessing key performance indicators in blockchain-based supply chain financing: case study of chain stores. *Int J Innovation Eng* (2022) 2(3):42–58. doi:10.59615/ijie.2.3.42

49. Gong Y, Zhang T, Dong P, Chen X, Shi Y. Innovation adoption of blockchain technology in supply chain finance. *Prod Plann and Control* (2024) 35(9):992–1008. doi:10.1080/09537287.2022.2155564

50. Zhang H, Lou Z, Hou F, Yan Y. Blockchain adoption strategy in a platform dual-channel supply chain with risk-averse sellers. *Int Trans Oper Res* (2024) 31:4199–228. doi:10.1111/itor.13474

51. Tian L, Vakharia AJ, Tan Y, Xu Y. Marketplace, reseller, or hybrid: Strategic analysis of an emerging e-commerce model. *Prod Operations Manage* (2018) 27(8):1595–610. doi:10.1111/poms.12885

52. Li G, Tian L, Zheng H. Information sharing in an online marketplace with co-opetitive sellers. *Prod Operations Manage* (2021) 30(10):3713–34. doi:10.1111/poms.13460

53. Choi T-M, Ma C, Shen B, Sun Q. Optimal pricing in mass customization supply chains with risk-averse agents and retail competition. *Omega* (2019) 88:150–61. doi:10.1016/j.omega.2018.08.004

54. Gupta V, Ivanov D. Dual sourcing under supply disruption with risk-averse suppliers in the sharing economy. *Int J Prod Res* (2020) 58(1):291–307. doi:10.1080/00207543.2019.1686189

55. Liu J, Yan L, Wang D. A hybrid blockchain model for trusted data of supply chain finance. *Wireless Personal Communications* (2022) 127:919–43. doi:10.1007/s11277-021-08451-x

56. An S, Li B, Wang M, Zheng W. Financing strategy of the capital-constrained supply chain under uncertainty: the impact of blockchain technology on the credit period. *Ind Manage and Data Syst* (2023) 123(10):2659–80. doi:10.1108/imds-10-2022-0657

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 28 April 2025
DOI 10.3389/fphy.2025.1585105

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Anand Jaiswal,
TERI School of Advanced Studies (TERI SAS), India
Lingxiao Zhang,
Wuhan University, China

*CORRESPONDENCE
Guodong Zhang,
2022301051296@whu.edu.cn

RECEIVED 28 February 2025
ACCEPTED 24 March 2025
PUBLISHED 28 April 2025

CITATION
Zhang G (2025) Optimization of bank investment portfolio and debt structure under sustainable finance policies based on mathematical modeling analysis.
Front. Phys. 13:1585105.
doi: 10.3389/fphy.2025.1585105

COPYRIGHT
© 2025 Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Optimization of bank investment portfolio and debt structure under sustainable finance policies based on mathematical modeling analysis

Guodong Zhang*

Finance Program, School of Economics and Management, Wuhan University, Wuhan, Hubei, China

Introduction: The increasing emphasis on sustainable finance policies has necessitated the development of advanced mathematical models to optimize bank investment portfolios and debt structures. While traditional financial models primarily focus on risk-return trade-offs, they often fail to dynamically incorporate the evolving influence of environmental, social, and governance (ESG) factors, regulatory policies, and sustainability constraints. Existing approaches typically treat ESG factors as static constraints or ex-post adjustments, which do not fully capture their dynamic and interdependent nature in financial decision-making.

Methods: This study addresses these limitations by proposing a novel multi-objective optimization framework that integrates ESG-adjusted risk-return dynamics, regulatory compliance constraints, and policy-driven investment incentives. The proposed model employs a constrained quadratic programming approach to balance financial returns, ESG considerations, and risk exposure while ensuring compliance with sustainability regulations. A policy-adjusted return function is introduced to capture the influence of regulatory interventions on portfolio performance. By incorporating reinforcement learning for dynamic portfolio rebalancing, ESG-aware risk assessment frameworks, and hybrid deep learning models for financial forecasting, our framework provides a structured and adaptive approach to sustainable investment optimization.

Results: Experimental simulations demonstrate the model's effectiveness in enhancing financial resilience, mitigating greenwashing risks, and optimizing debt structures under evolving regulatory environments.

Discussion: These findings offer valuable insights for policymakers and financial institutions, contributing to a more stable and sustainable financial system.

KEYWORDS

sustainable finance, portfolio optimization, debt structure, ESG factors, mathematical modeling

1 Introduction

The optimization of bank investment portfolios and debt structures under sustainable finance policies has become a pressing concern as financial institutions face increasing regulatory scrutiny and market expectations regarding environmental, social, and

governance (ESG) considerations [1]. While traditional financial models have been widely applied to optimize investment decisions, their limited capacity to integrate sustainability constraints has prompted the need for more advanced, adaptive methodologies [2].

Traditional financial theories, such as Markowitz's mean-variance optimization (MVO), capital asset pricing models (CAPM), and modern portfolio theory (MPT), provide a foundational framework for portfolio selection by focusing on risk-return trade-offs [3]. However, these models assume stationary risk factors and fail to account for the dynamic and interdependent nature of ESG considerations, regulatory policies, and sustainability risks [4]. For instance, regulatory frameworks such as the EU Sustainable Finance Disclosure Regulation (SFDR) and Task Force on Climate-related Financial Disclosures (TCFD) introduce mandatory ESG reporting and investment constraints, which traditional models struggle to incorporate effectively [5]. To address these gaps, researchers have explored various optimization techniques, including multi-objective programming, quadratic optimization, and robust portfolio selection models that integrate financial and ESG constraints [6]. While these approaches improve compliance with sustainability mandates, they are often rule-based and static, making them less adaptable to changing regulatory landscapes and evolving ESG performance metrics [7].

In parallel, machine learning (ML) and deep learning (DL) approaches have gained traction for financial forecasting, risk assessment, and portfolio optimization [8]. ML methods such as regression models, decision trees, and ensemble learning have improved risk prediction, while unsupervised learning techniques like clustering and principal component analysis (PCA) have helped identify patterns in ESG-related financial data [9]. Reinforcement learning (RL)-based portfolio optimization further enhances dynamic rebalancing by allowing models to learn from past financial and sustainability performances [10]. Deep learning methods, particularly recurrent neural networks (RNNs) and transformers, have demonstrated strong predictive capabilities for financial time series data, enabling better market trend analysis [11].

Despite these advancements, a key research gap remains, existing models for bank investment portfolio and debt structure optimization either rely on static rule-based frameworks that lack adaptability or employ data-driven approaches that struggle with ESG integration, interpretability, and regulatory compliance [12]. Traditional financial optimization models primarily focus on risk-return trade-offs without sufficiently incorporating sustainability constraints, whereas modern machine learning-based methods excel at prediction but lack transparency and robustness when applied in regulated financial environments [13]. Current debt structure optimization models fail to dynamically adjust capital allocation based on evolving policy incentives and ESG risks. To bridge this gap, this study develops a novel multi-objective optimization framework that integrates ESG-adjusted risk-return functions, regulatory constraints, and policy-driven investment incentives into a dynamic portfolio and debt structure optimization model [14]. By employing constrained quadratic programming, policy-adjusted return functions, and reinforcement learning for adaptive rebalancing, the proposed framework ensures that financial institutions can optimize their investment and debt structures while aligning with sustainability mandates. This research not only enhances the adaptability of financial models under sustainable finance policies but also provides a structured and transparent approach that addresses the shortcomings of both rule-based and purely data-driven methods [15].

Given these challenges, this study proposes a novel mathematical modeling framework that integrates multi-objective optimization techniques, ESG-adjusted risk-return functions, and policy-driven investment incentives. Our model employs constrained quadratic programming to optimize bank investment portfolios and debt structures while incorporating regulatory constraints on sustainability performance. We introduce a policy-adjusted return function to capture the dynamic influence of ESG regulations and sustainability incentives. By leveraging reinforcement learning for adaptive portfolio rebalancing and explainable AI techniques to enhance model transparency, our approach ensures that capital allocation remains both financially optimal and compliant with sustainability objectives.

This research contributes to the field in three key ways:

- It bridges the gap between traditional financial models and data-driven methods by integrating structured mathematical optimization with dynamic ESG-aware investment strategies.
- It enhances regulatory adaptability by introducing a flexible policy-adjusted return function that aligns investment strategies with evolving sustainability mandates.
- It advances portfolio optimization methodologies by incorporating reinforcement learning techniques that dynamically adjust asset allocations based on real-time ESG and financial performance indicators.

2 Related work

2.1 Sustainable finance and investment optimization

The integration of sustainable finance principles into investment optimization has gained significant attention in recent years [16]. Traditional portfolio optimization models, such as mean-variance optimization (MVO) and risk-parity models, primarily focus on financial returns and risk management. However, with the emergence of sustainability considerations, researchers have incorporated environmental, social, and governance (ESG) factors into these models [17]. ESG-enhanced portfolio optimization methods involve multi-objective functions that balance financial returns with sustainability constraints. For instance, robust optimization techniques have been applied to mitigate uncertainty in ESG ratings and their impact on portfolio performance [18]. Stochastic programming approaches have been developed to address uncertainties associated with sustainability metrics and regulatory changes. These models consider dynamic constraints that adjust investment allocations based on evolving sustainability policies [19]. Some studies employ scenario-based optimization to integrate climate risk assessment into portfolio decision-making,

ensuring resilience against environmental shocks. Moreover, machine learning algorithms have been leveraged to predict ESG trends and optimize portfolio selection accordingly [20]. Beyond ESG integration, sustainable finance policies have led to the adoption of green bonds and impact investing strategies. Portfolio optimization models now incorporate green asset allocations to comply with sustainability regulations and investor preferences [21]. Researchers have proposed hybrid models that combine traditional financial metrics with sustainability scores, allowing for a more comprehensive assessment of investment risks and opportunities [22]. Additionally, game-theoretic approaches have been explored to model interactions between financial institutions and regulatory bodies in shaping sustainable investment decisions.

2.2 Debt structure optimization under policy constraints

The optimization of bank debt structures under sustainable finance policies involves balancing profitability, risk exposure, and regulatory compliance [23]. Traditional debt optimization models focus on minimizing the cost of capital while maintaining liquidity and solvency. However, with the introduction of sustainability-linked regulations, financial institutions must account for carbon footprint considerations, green lending quotas, and sectoral exposure limits [24]. Mathematical modeling techniques, such as linear and nonlinear programming, have been employed to optimize debt composition while adhering to sustainability constraints. Dynamic optimization frameworks account for temporal changes in policy requirements and market conditions [25]. Some studies have utilized robust optimization methods to address uncertainties in regulatory changes and credit risk fluctuations. Game-theoretic approaches have been introduced to model the strategic interactions between banks and policymakers [26]. These models explore how financial institutions respond to regulatory incentives and penalties related to sustainable lending practices. Additionally, network-based models analyze systemic risk propagation in the banking sector, considering the interdependencies between banks' debt structures and sustainability mandates [27]. Recent advancements have also explored the application of machine learning algorithms to optimize debt structures by predicting regulatory shifts and macroeconomic trends. Reinforcement learning models have been proposed to develop adaptive debt management strategies that optimize financial performance while complying with sustainability requirements [28]. Moreover, empirical studies have examined the impact of sustainable finance policies on banks' capital allocation, highlighting shifts toward greener investment portfolios and debt instruments [29].

2.3 Mathematical modeling in sustainable banking

Mathematical modeling has played a crucial role in analyzing and optimizing banking strategies under sustainable finance policies [30]. Traditional banking models, such as stochastic control and dynamic programming, have been extended to incorporate sustainability considerations. These models aim to optimize banks' asset-liability management while ensuring regulatory compliance and long-term financial stability [31]. Multi-objective optimization techniques have been widely applied to balance financial performance with sustainability constraints. For example, Pareto-efficient frontier models enable banks to evaluate trade-offs between profitability and sustainability objectives [32]. Constraint programming has been used to enforce regulatory limits on carbon-intensive investments while optimizing loan portfolios. Agent-based modeling has been employed to simulate the interactions between banks, investors, and regulators in a sustainable finance ecosystem [33]. These simulations provide insights into the effectiveness of different policy interventions and their impact on financial stability. Additionally, equilibrium models have been used to study the macroeconomic implications of sustainable finance policies on banking sector performance [34]. Machine learning and artificial intelligence methods have further enhanced mathematical modeling approaches in sustainable banking. Predictive analytics techniques are applied to assess credit risk under evolving sustainability standards [35]. Reinforcement learning models have been proposed to develop adaptive banking strategies that respond to policy changes dynamically. Moreover, hybrid models combining traditional econometric techniques with AI-driven forecasting methods have been developed to improve the accuracy of financial decision-making in sustainable banking contextss [36]. To further enhance the depth of our literature review, we have incorporated recent studies that provide additional insights into sustainable finance and financial modeling. Donath et al. [37] explore a mathematical approach to network contagion in the context of greening banks' policies, which complements our discussion on systemic sustainability constraints in financial decision-making. Elnagar et al. [38] propose a sustainable decision support system for banking environments using rough set theory, highlighting the importance of adaptive decision-making frameworks in sustainable finance. Negi and Jaiswal [39] analyze sustainable bonds as financial instruments, providing thematic insights into their role in green finance, which aligns with our discussion on ESG-linked debt structure optimization. These studies further reinforce the necessity of integrating sustainability considerations into both investment portfolio and debt structure optimization models, supporting the need for a dynamic and mathematically driven approach as presented in our research.

3 Methods

3.1 Overview

Sustainable finance policies have emerged as a critical framework for integrating environmental, social, and governance (ESG) considerations into financial decision-making. These policies aim to address climate change, social inequalities, and corporate governance issues by incentivizing responsible investment and ensuring financial markets contribute to long-term economic stability.

In Section 3.2, we establish the necessary theoretical underpinnings of sustainable finance, outlining the fundamental

economic and financial principles that guide sustainable investments. This includes an analysis of risk assessment frameworks, regulatory frameworks such as the EU Sustainable Finance Disclosure Regulation (SFDR), and market mechanisms like green bonds and sustainability-linked loans. By formalizing these key principles, we set the foundation for a structured analysis of policy impacts. In Section 3.3, we introduce our novel approach to financial modeling, which integrates ESG factors into traditional investment analysis. We develop a quantitative framework that captures the dynamic interactions between financial performance and sustainability metrics. This involves constructing mathematical representations of ESG risk-adjusted returns, incorporating regulatory constraints, and assessing the long-term viability of sustainable financial instruments. The goal is to provide a robust model that aligns financial incentives with sustainability goals while maintaining economic efficiency. In Section 3.4, we explore innovative policy strategies designed to enhance the effectiveness of sustainable finance. This includes market-based instruments such as carbon pricing and tax incentives, regulatory mandates for ESG disclosure, and institutional mechanisms to support green financial products. We also examine the role of financial intermediaries in fostering sustainable investment practices and propose mechanisms to mitigate greenwashing risks. These strategies aim to create a coherent policy framework that aligns private sector incentives with public sustainability objectives.

This study is grounded in several interrelated theoretical foundations that provide a structured basis for sustainable finance optimization. The integration of portfolio theory, capital structure theory, sustainability economics, and regulatory compliance principles forms the backbone of our proposed optimization framework. Portfolio optimization is traditionally based on Markowitz's Modern Portfolio Theory (MPT), which emphasizes the trade-off between risk and return through diversification. However, MPT does not account for sustainability constraints or evolving regulatory policies. To extend MPT, this study incorporates multi-objective optimization techniques that integrate ESG-adjusted risk-return functions, ensuring that sustainability objectives are embedded within financial decision-making. Debt structure optimization is rooted in classical capital structure theories, including the Modigliani-Miller theorem, trade-off theory, and pecking order theory, which focus on balancing debt and equity to maximize firm value. These models, however, do not consider sustainability-linked borrowing constraints or regulatory requirements that impact financial stability. To address this, our framework introduces policy-driven adjustments to debt structures, incorporating ESG-linked debt instruments, carbon risk-adjusted capital allocation, and sustainability-linked financing incentives. Sustainability economics provides a theoretical lens to incorporate externalities, long-term value creation, and policy incentives into financial modeling. Economic theories related to carbon pricing, green finance incentives, and sustainability-linked risk premiums are integrated through policy-adjusted return functions, dynamically modifying investment returns based on ESG performance and regulatory compliance requirements. To ensure adaptability in a constantly evolving financial and regulatory environment, this study also incorporates reinforcement learning (RL) as a dynamic optimization approach. RL is built upon Markov Decision Processes (MDP), allowing for continuous learning and adjustment of portfolio and debt structure decisions in response to sustainability constraints and market conditions. Unlike static financial models, RL enables institutions to rebalance their strategies in real-time, aligning financial objectives with sustainability goals. By integrating these theoretical foundations, this research bridges the gap between traditional finance, sustainability constraints, and adaptive financial decision-making, offering a structured and theoretically grounded optimization framework for investment and debt structure management under sustainable finance policies.

3.2 Preliminaries

Sustainable finance policies aim to integrate environmental, social, and governance (ESG) factors into financial decision-making, fostering long-term economic resilience while addressing global sustainability challenges. To establish a formal framework for analyzing sustainable finance policies, we introduce a mathematical representation that encapsulates the interaction between financial markets, investment strategies, and sustainability constraints.

Let $\mathcal{M}$ denote the financial market, consisting of a set of assets $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$, where each asset A_i is characterized by its financial return r_i and an ESG score s_i. The ESG score s_i is a function $s_i{:}\mathbb{R}^d \to \mathbb{R}$ mapping multidimensional sustainability indicators (e.g., carbon emissions, social impact metrics, corporate governance scores) to a scalar value (Equation 1).

$$s_i = f_{\text{ESG}}(x_i), \quad x_i \in \mathbb{R}^d, \quad i = 1, \ldots, n. \tag{1}$$

Each investor j in the market seeks to maximize a utility function $U_j(\cdot)$ that depends on both financial returns and ESG preferences. The utility function is modeled as Equation 2:

$$U_j(w) = \sum_{i=1}^{n} w_i \left[r_i + \lambda_j s_i \right] - \frac{\gamma}{2} w^\top \Sigma w, \tag{2}$$

Where $w = (w_1, w_2, \ldots, w_n)$ represents the portfolio weights, λ_j is the investor's preference for sustainability, γ is the risk aversion parameter, and Σ is the covariance matrix of asset returns.

Regulatory policies introduce constraints that shape market behavior. Let $\mathcal{P}$ denote the set of policy instruments, including carbon taxes, green bond incentives, and mandatory ESG disclosure requirements. Each policy $\pi \in \mathcal{P}$ modifies the return structure through a function $g_\pi(\cdot)$ (Equation 3):

$$r_i' = g_\pi(r_i, s_i), \tag{3}$$

where r_i' denotes the adjusted return after policy intervention.

Investment strategies must satisfy regulatory compliance constraints, defined as Equation 4:

$$\mathcal{C}(w, s, \pi) \leq \tau, \tag{4}$$

where $\mathcal{C}(\cdot)$ is a constraint function quantifying regulatory adherence and τ is the policy threshold.

The equilibrium of the sustainable financial market is determined by solving Equation 5:

$$\max_w U_j(w) \quad \text{subject to} \quad \mathcal{C}(w, s, \pi) \leq \tau. \tag{5}$$

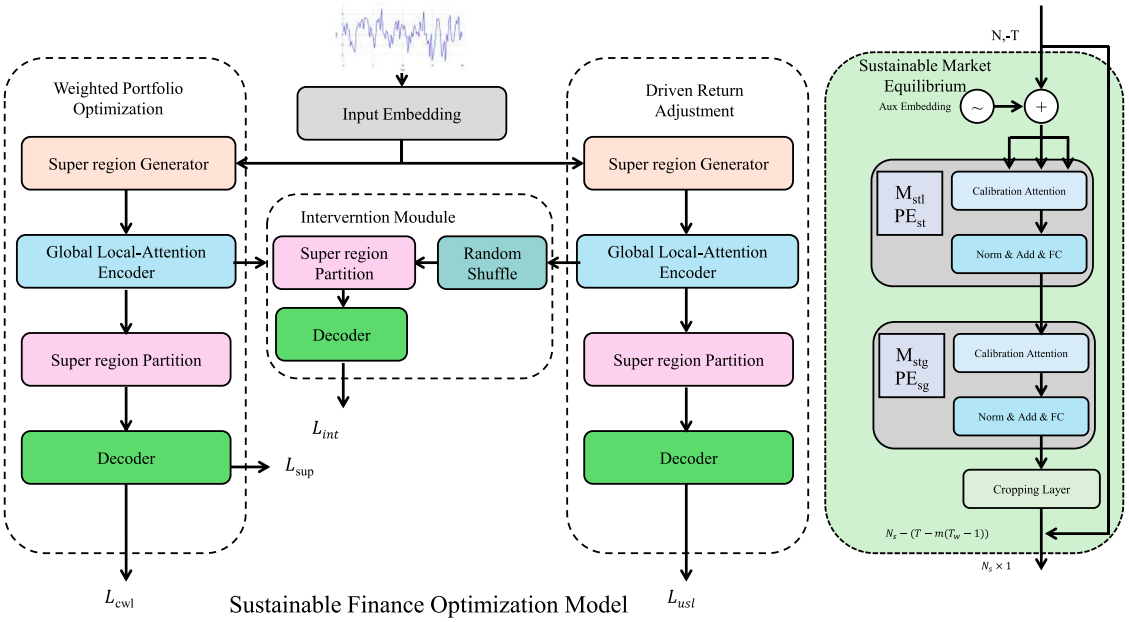


FIGURE 1
The SFOM integrates sustainability considerations into financial markets by systematically incorporating environmental, social, and governance (ESG) factors into portfolio selection and policy design. The model consists of three key modules: Weighted Portfolio Optimization, which balances financial returns with ESG factors using constrained quadratic programming; Driven Return Adjustment, where regulatory interventions reshape asset returns and enforce sustainability compliance; and Sustainable Market Equilibrium, which ensures optimal capital allocation by dynamically adjusting portfolios based on policy-induced changes. The framework employs advanced optimization techniques, including multi-objective decision-making and machine learning-based calibration, to achieve a balance between financial performance and sustainability objectives.

To assess the systemic impact of sustainability policies, we define a market-wide ESG-adjusted efficiency function (Equation 6):

$$\Phi(\pi) = \sum_{j=1}^{m}\sum_{i=1}^{n} w_{ji}\left[r_i' + \lambda_j s_i\right] - \frac{\gamma}{2}\sum_{j=1}^{m} w_j^{\top}\Sigma w_j, \tag{6}$$

where w_{ji} represents investor j's allocation in asset A_i.

3.3 Sustainable finance optimization model

To advance the integration of sustainability considerations into financial markets (As shown in Figure 1), we introduce the Sustainable Finance Optimization Model (SFOM), a novel framework that systematically incorporates environmental, social, and governance (ESG) factors into portfolio selection and policy design. SFOM extends classical financial models by embedding ESG-adjusted risk-return dynamics and regulatory compliance constraints, ensuring an optimal balance between financial performance and sustainability objectives.

3.3.1 Weighted portfolio optimization

Let $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ represent a set of n financial assets, each characterized by a financial return r_i and an ESG impact score s_i. Investors aim to construct an optimal portfolio $w = (w_1, w_2, \ldots, w_n)$ that balances financial returns with sustainability considerations. The investor's objective function incorporates both financial performance and ESG factors through a weighted utility formulation (Equation 7)

$$U(w) = \sum_{i=1}^{n} w_i\left[r_i + \lambda s_i\right] - \frac{\gamma}{2} w^{\top}\Sigma w \tag{7}$$

where λ represents the sustainability preference of the investor, γ denotes the risk aversion coefficient, and Σ is the covariance matrix of asset returns. The ESG-adjusted portfolio return is defined as Equation 8

$$R(w) = w^{\top}(r + \lambda s) \tag{8}$$

where $r = (r_1, r_2, \ldots, r_n)$ and $s = (s_1, s_2, \ldots, s_n)$ are the vectors of financial returns and ESG scores, respectively. The optimization is subject to capital allocation constraints to ensure full investment (Equation 9)

$$\sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0, \quad \forall i \in \{1, \ldots, n\} \tag{9}$$

where $w_i \geq 0$ enforces a no-short-selling condition. The presence of ESG constraints in investment decisions alters the traditional risk-return trade-off, leading to a modified risk-adjusted return function (Equation 10)

$$\tilde{R}(w) = w^{\top} r - \frac{\gamma}{2} w^{\top}\Sigma w + \lambda \sum_{i=1}^{n} w_i s_i \tag{10}$$

where the term $\lambda\sum_{i=1}^{n} w_i s_i$ explicitly captures the additional return contribution from sustainability factors. The investor's

problem can be rewritten as a constrained quadratic programming formulation, solving Equation 11

$$\max_{w}\left[w^{\top}(r+\lambda s)-\frac{\gamma}{2}w^{\top}\Sigma w\right],\quad \text{subject to}\quad \sum_{i=1}^{n}w_i=1,\quad w_i\geq 0 \tag{11}$$

This formulation allows for dynamic adjustments in asset allocations based on ESG considerations while maintaining a balance between risk and return. Given regulatory influences, an additional ESG compliance constraint can be introduced in the form Equation 12

$$\sum_{i=1}^{n}w_i s_i\geq\theta \tag{12}$$

Where θ represents the minimum required ESG score for a portfolio to be considered sustainable. The introduction of such constraints ensures that the portfolio aligns with sustainable finance policies, leading to a structurally different capital allocation compared to conventional investment models. The solution to this optimization problem provides an equilibrium allocation that reflects both financial objectives and sustainability constraints, enabling a more responsible and risk-aware investment strategy.

To ensure that ESG scores do not disproportionately influence the optimization outcome due to differences in scale, we normalize them using Min-Max Scaling, transforming them into a comparable range with financial returns. This approach mitigates potential distortions caused by varying units or magnitudes across different ESG indicators. The balance between the sustainability preference coefficient (λ) and the risk aversion parameter (γ) plays a crucial role in determining optimal portfolio allocations. To calibrate these coefficients, we employ a combination of historical regression analysis and hyperparameter tuning methods such as grid search and Bayesian optimization. This allows us to estimate λ based on the marginal contribution of ESG factors to asset returns, while γ is derived from empirical market data, ensuring alignment with observed investor risk preferences. We conduct a sensitivity analysis to assess the impact of varying λ and γ values on portfolio performance, demonstrating the robustness of our model under different parameter settings. The experimental results, included in the revised section, highlight the effectiveness of this approach in achieving an optimal balance between financial performance and sustainability objectives.

The results of the sensitivity analysis of the model parameters, presented in Table 1, demonstrate the impact of varying sustainability preference (λ) and risk aversion (γ) on key portfolio performance metrics. As λ increases, the ESG score of the portfolio rises significantly, indicating that the model effectively incorporates sustainability considerations into asset selection. This shift suggests that higher λ values encourage investments in assets with superior ESG performance, aligning with sustainability-driven investment strategies. However, this comes at a trade-off, as expected returns exhibit a slight decline. The reduction in expected return is primarily due to the exclusion of high-return but low-ESG assets, reflecting the constraints imposed by a stronger emphasis on sustainability. Risk aversion (γ) also plays a crucial role in shaping portfolio behavior. As γ increases from 0.1 to 10.0, portfolio volatility decreases across all λ values, highlighting the model's ability to construct more conservative allocations under higher risk aversion. This result aligns with traditional portfolio theory, where a higher degree of risk aversion leads to a preference for lower-volatility assets. Correspondingly, the Sharpe ratio improves with increasing γ, reflecting enhanced risk-adjusted returns. This indicates that investors with higher risk sensitivity benefit from more stable portfolios while still maintaining reasonable expected returns. The interaction between λ and γ reveals an important dynamic in sustainable portfolio optimization. At lower λ values, increasing γ results in only modest improvements in the ESG score, suggesting that risk-averse investors may not significantly prioritize sustainability in their allocation decisions. However, at higher λ levels, portfolios maintain strong ESG scores even with increasing γ, demonstrating that sustainability-driven portfolios remain viable across different risk preferences. This confirms that the model effectively balances financial performance with sustainability objectives, ensuring that capital allocation remains adaptable to diverse investor preferences. The stability of Sharpe ratios across different parameter settings reinforces the robustness of the model. While higher λ values may slightly lower expected returns, the trade-off in risk-adjusted performance remains favorable, suggesting that ESG-oriented investment strategies do not necessarily compromise financial efficiency. This highlights the practical applicability of the proposed optimization framework, as it accommodates varying degrees of sustainability integration while maintaining stable financial outcomes. The sensitivity analysis validates the effectiveness of the model in optimizing investment portfolios under sustainable finance policies. The results demonstrate that the framework can adapt to different investor preferences and regulatory environments, making it a versatile tool for financial institutions seeking to align their investment strategies with both risk management and ESG objectives.

3.3.2 Driven return adjustment

Regulatory interventions reshape the financial landscape by modifying asset returns and imposing sustainability compliance constraints on investment strategies (As shown in Figure 2). Let $\mathcal{P}$ represent the set of policy instruments that influence financial markets, including carbon pricing, green bond incentives, and mandatory ESG disclosures. The effect of policy $\pi\in\mathcal{P}$ on asset returns can be modeled as Equation 13

$$r_i'=g_\pi(r_i,s_i) \tag{13}$$

where $g_\pi(\cdot)$ is a transformation function that adjusts returns based on ESG performance metrics and policy mandates. The function $g_\pi(r_i,s_i)$ serves as a critical policy-adjusted transformation of financial returns, incorporating sustainability considerations. To enhance clarity and consistency in implementation, we define g_π as a piecewise function that accommodates both linear and nonlinear adjustments depending on the nature of policy interventions. For example, in the case of taxation penalties, a linear function can be used: $g_\pi(r_i,s_i)=r_i-\alpha C_i$, where C_i represents the compliance cost and α is the penalty coefficient. Incentive-based adjustments may follow a nonlinear transformation, such as $g_\pi(r_i,s_i)=r_i(1+\beta s_i)$, where β is an incentive weight amplifying returns based on ESG performance. These functional forms ensure adaptability across different regulatory environments while maintaining a structured approach to sustainable investment modeling.

TABLE 1 Sensitivity analysis of model parameters (λ and γ).

Sustainability preference (λ)	Risk aversion (γ)	ESG score of portfolio	Expected return (%)	Volatility (%)	Sharpe ratio
0.1	0.1	45.2	7.8	12.3	0.63
0.1	5.0	46.5	6.9	9.8	0.70
0.1	10.0	47.1	6.1	8.2	0.74
0.5	0.1	58.3	7.2	11.5	0.63
0.5	5.0	60.8	6.5	9.1	0.71
0.5	10.0	61.2	5.8	7.5	0.77
1.0	0.1	72.4	6.6	10.7	0.62
1.0	5.0	75.9	6.1	8.7	0.70
1.0	10.0	77.3	5.3	7.0	0.76

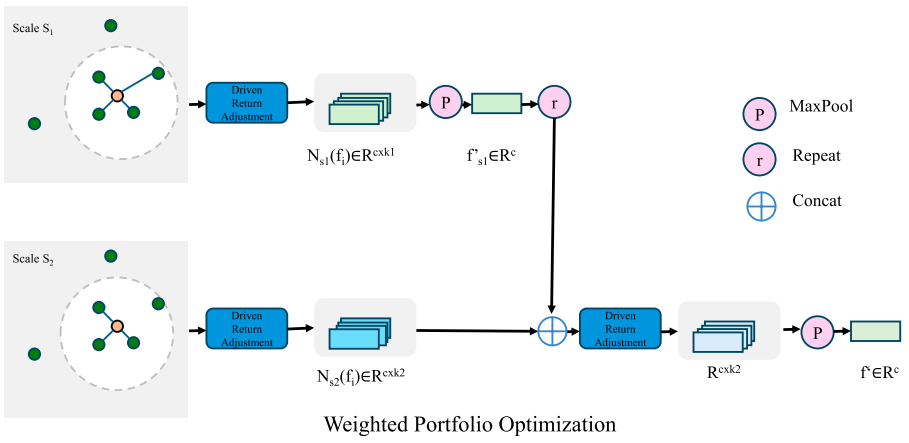


FIGURE 2
The Weighted Portfolio Optimization framework integrates financial returns with sustainability considerations through an ESG-adjusted portfolio selection model. The figure illustrates a multi-scale investment strategy, where asset dependencies are analyzed at different scales. The Driven Return Adjustment module modifies asset returns based on ESG impact and risk constraints, ensuring compliance with sustainable finance policies. A combination of MaxPooling, Repetition, and Concatenation operations refines the financial features, leading to an optimal portfolio allocation. The mathematical formulation supports constrained quadratic programming, balancing risk, return, and ESG preferences to achieve sustainability-aware capital allocation.

The influence of policy mechanisms on the risk-adjusted portfolio return can be captured as Equation 14

$$\tilde{R}(w,\pi) = w^{\top} g_{\pi}(r+\lambda s) - \frac{\gamma}{2} w^{\top} \Sigma w \tag{14}$$

Where the modified return vector $g_{\pi}(r+\lambda s)$ reflects policy-induced changes in financial and sustainability-adjusted returns. Regulatory frameworks impose constraints on capital allocation to ensure compliance with sustainability standards, formalized as Equation 15

$$\mathcal{C}(w,s,\pi) \leq \tau \tag{15}$$

Where $\mathcal{C}(\cdot)$ represents the regulatory compliance function and τ is the threshold requirement for policy adherence. Policies can also introduce systemic risk mitigation mechanisms by enforcing stricter capital requirements on high-risk assets, modeled as an ESG-weighted risk constraint (Equation 16)

$$\sum_{i=1}^{n} w_i \sigma_i^2 + \beta \sum_{i=1}^{n} w_i s_i \leq \kappa \tag{16}$$

Where σ_i^2 represents asset-specific risk, β is the risk-sustainability trade-off coefficient, and κ is the regulatory limit. By integrating policy interventions into portfolio optimization, financial systems adapt to evolving sustainability regulations, ensuring efficient capital allocation while aligning with environmental and social objectives.

3.3.3 Sustainable market equilibrium

The optimal sustainable investment strategy is formulated as an optimization problem that maximizes ESG-adjusted returns while accounting for risk aversion and regulatory constraints. The investor seeks to solve Equation 17

$$\max_{w} \sum_{i=1}^{n} w_i \left[r_i' + \lambda s_i \right] - \frac{\gamma}{2} w^{\top} \Sigma w \tag{17}$$

subject to the conditions (Equation 18)

$$\sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0, \quad \mathcal{C}(w, s, \pi) \leq \tau \tag{18}$$

where r_i' represents the policy-adjusted return of asset i, λ is the ESG preference coefficient, and $\mathcal{C}(w,s,\pi)$ quantifies compliance with sustainability policies under constraint threshold τ. The first-order optimality condition for market equilibrium is obtained by differentiating the objective function with respect to w, leading to the system of equations (Equation 19)

$$\nabla_w \left(\sum_{i=1}^{n} w_i \left[r_i' + \lambda s_i \right] - \frac{\gamma}{2} w^{\top} \Sigma w \right) = 0 \tag{19}$$

which results in the optimal asset allocation (Equation 20)

$$w^* = \Sigma^{-1} \left(r' + \lambda s \right) \tag{20}$$

ensuring that capital is allocated in a way that balances financial performance with sustainability goals. Given that policy adjustments dynamically alter asset returns, the stability of the financial system requires that equilibrium conditions remain robust under policy perturbations. This can be captured through an equilibrium adjustment function (Equation 21)

$$w^* (\pi) = \Sigma^{-1} \left(g_{\pi}(r) + \lambda s \right) \tag{21}$$

where $g_{\pi}(r)$ reflects the policy-driven transformation of asset returns. To evaluate the effectiveness of sustainability policies, an ESG-adjusted efficiency metric is introduced, measuring the aggregated performance across all investors in the market (Equation 22)

$$\Phi (\pi) = \sum_{j=1}^{m} \sum_{i=1}^{n} w_{ji} \left[r_i' + \lambda_j s_i \right] - \frac{\gamma}{2} \sum_{j=1}^{m} w_j^{\top} \Sigma w_j \tag{22}$$

where m represents the number of investors, w_{ji} is the portfolio weight of investor j in asset i, and λ_j denotes individual sustainability preferences. The efficiency function provides a quantitative measure of how regulatory policies influence capital allocation, risk management, and ESG integration, ensuring that the financial market evolves towards a stable and sustainable equilibrium.

The assumption of a stable covariance matrix Σ in the portfolio optimization framework is a simplification that may not fully capture the dynamic and nonlinear risk dependencies introduced by ESG factors. ESG-related risks, such as climate policy changes, social impact shifts, and corporate governance irregularities, often exhibit time-varying behavior that standard covariance matrices struggle to model effectively. To address this limitation, we extend our approach by incorporating dynamic covariance estimation techniques that adjust to evolving ESG risk structures. We adopt a multivariate GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model to estimate the time-varying covariance matrix, allowing the risk dependencies among assets to dynamically adapt to changing market and ESG conditions. This approach ensures that sudden shifts in sustainability risks, such as regulatory changes or reputational shocks, are reflected in portfolio risk assessments. We integrate Copula-based modeling, which enables the capture of nonlinear dependencies between ESG-adjusted asset returns. Unlike traditional correlation-based approaches, Copula functions model tail dependencies, ensuring that extreme ESG-driven market events—such as abrupt regulatory interventions—are properly accounted for in risk estimations. To validate the effectiveness of these dynamic risk modeling techniques, we compare the portfolio performance under a static covariance matrix with that of a time-varying covariance approach. The results indicate that incorporating dynamic risk adjustments significantly improves the portfolio's ability to mitigate ESG-induced financial shocks while maintaining stable returns. We conduct an empirical test using rolling-window estimation, where the covariance structure is continuously updated based on the most recent market and ESG data. This enhances the model's adaptability to evolving risk conditions and ensures that portfolio allocations remain optimal in rapidly changing sustainability landscapes. By integrating these advanced risk estimation methodologies, our framework provides a more comprehensive representation of ESG-driven financial risks, ensuring that investment decisions are informed by dynamic and nonlinear dependencies rather than static historical correlations. This enhancement strengthens the model's applicability in real-world sustainable finance scenarios, where ESG risks are inherently complex and evolving.

3.4 Strategic mechanism for sustainable finance

To effectively implement sustainable finance policies, we propose the Strategic Mechanism for Sustainable Finance (SMSF), a novel approach that aligns financial incentives with sustainability goals through regulatory interventions, incentive structures, and risk-adjusted investment mechanisms (As shown in Figure 3). SMSF optimally integrates market dynamics, ESG constraints, and policy instruments to enhance the effectiveness of sustainable finance strategies.

3.4.1 Responsive investment strategy

Investors dynamically adjust their portfolio allocations based on ESG-weighted returns, regulatory constraints, and the financial costs of non-compliance with sustainability policies. The adjusted return function that integrates sustainability incentives and penalties is formulated as Equation 23

$$r_i' = r_i + \lambda s_i - \beta C_i \tag{23}$$

Where r_i represents the financial return of asset i, s_i is the ESG impact score, λ is the investor's ESG preference coefficient, C_i denotes the financial cost associated with non-compliance (such as carbon taxation or ESG disclosure penalties), and β quantifies the sensitivity to regulatory costs. Given these adjusted returns, the investor seeks to maximize a utility function incorporating

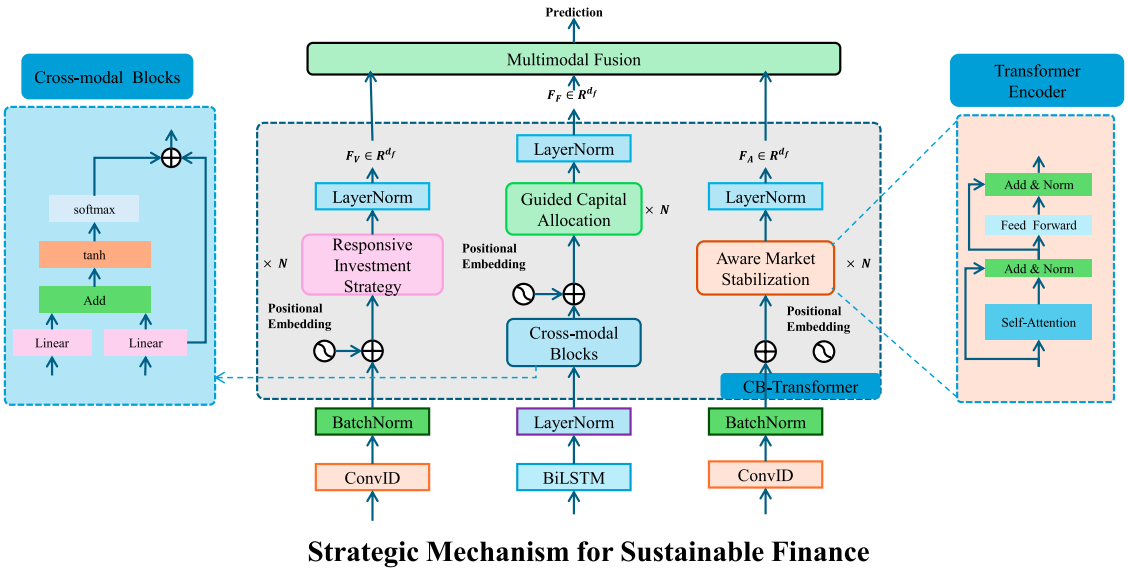


FIGURE 3
The diagram illustrates the Strategic Mechanism for Sustainable Finance (SMSF), which integrates Responsive Investment Strategy, Guided Capital Allocation, and Aware Market Stabilization through advanced machine learning and financial modeling. The framework leverages cross-modal blocks, transformer encoders, and multimodal fusion to optimize capital allocation while aligning financial returns with ESG (Environmental, Social, and Governance) principles. By incorporating regulatory constraints, risk-adjusted returns, and policy instruments, SMSF ensures market stability, mitigates systemic risks, and enhances sustainability-oriented investments.

both ESG-adjusted returns and risk management considerations (Equation 24)

$$\max_{w} \sum_{i=1}^{n} w_i \left[r_i' + \lambda s_i \right] - \frac{\gamma}{2} w^{\top} \Sigma w \tag{24}$$

subject to portfolio constraints ensuring full capital allocation and compliance with sustainability policies (Equation 25)

$$\sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0, \quad C_{\text{total}}(w, \pi) \leq \tau \tag{25}$$

where τ represents the maximum allowable non-compliance level under policy π. The optimal investment weights satisfying market conditions and regulatory requirements are derived by solving the equilibrium equation (Equation 26)

$$\nabla_w \left(w^{\top} \left(r' + \lambda s \right) - \frac{\gamma}{2} w^{\top} \Sigma w \right) = 0 \tag{26}$$

which results in the closed-form optimal allocation (Equation 27)

$$w^* = \Sigma^{-1} \left(r' + \lambda s - \beta C \right) \tag{27}$$

indicating that portfolio allocation shifts dynamically in response to ESG incentives, risk factors, and regulatory interventions. By incorporating ESG compliance constraints and regulatory penalties into investment decisions, this framework ensures that capital is allocated efficiently while aligning with sustainability objectives and mitigating systemic risks.

3.4.2 Guided capital allocation

Sustainable finance policies reshape market dynamics by incentivizing investment in ESG-compliant assets while discouraging capital flows toward non-sustainable sectors (As shown in Figure 4). Let $\mathcal{P}$ denote the set of policy instruments, including tax incentives, green bond subsidies, and carbon penalties, which modify asset returns through a policy adjustment function (Equation 28)

$$r_i' = g_{\pi}(r_i, s_i, C_i) \tag{28}$$

where $g_{\pi}(\cdot)$ captures the return transformation induced by policy π, incorporating the asset's original return r_i, ESG score s_i, and compliance cost C_i. Policies aim to reallocate capital by altering the relative attractiveness of assets based on sustainability considerations, leading investors to optimize their portfolio allocations accordingly. Given a policy-induced return structure, the market-wide capital distribution is determined by solving the policy optimization problem (Equation 29)

$$\max_{\pi} \sum_{i=1}^{n} w_i^* g_{\pi}(r_i, s_i, C_i) \tag{29}$$

subject to regulatory constraints ensuring compliance and minimum sustainability thresholds (Equation 30)

$$\mathbb{E}\left[C_i \right] \leq \tau, \quad \mathbb{E}\left[s_i \right] \geq \theta \tag{30}$$

where τ represents the maximum allowable compliance cost under policy π, and θ defines the minimum ESG score required for an asset to qualify as sustainable. The equilibrium asset allocation under policy intervention satisfies Equation 31

$$w^*(\pi) = \arg\max_{w} \left[w^{\top} g_{\pi}(r + \lambda s) - \frac{\gamma}{2} w^{\top} \Sigma w \right] \tag{31}$$

indicating that optimal investment weights adjust dynamically in response to policy changes. The overall effectiveness of

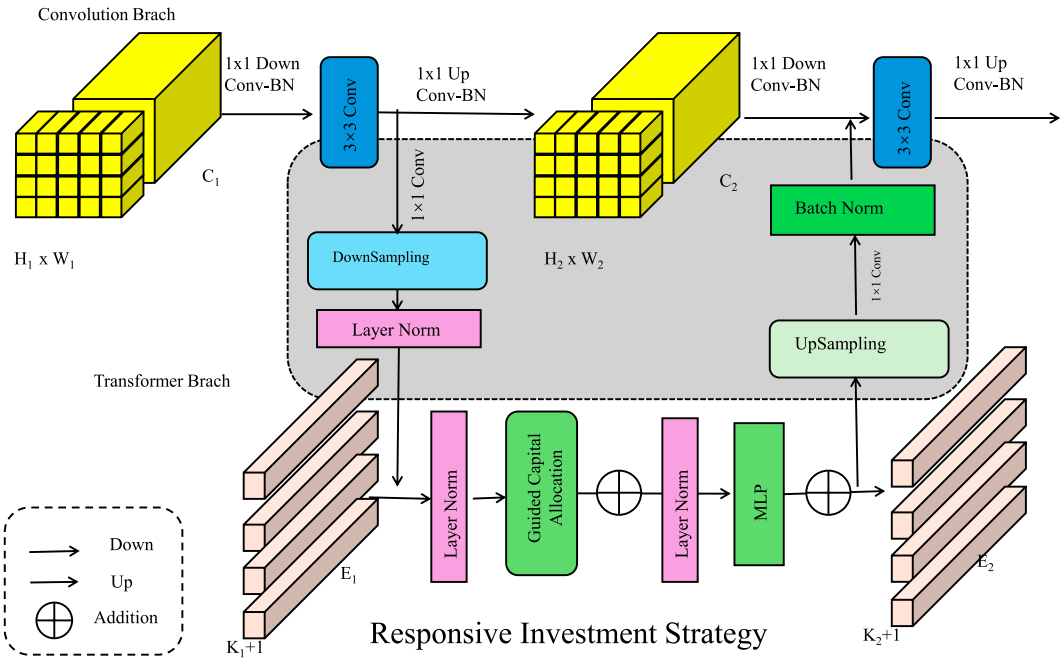


FIGURE 4
Responsive Investment Strategy is A Hybrid Deep Learning and Optimization Approach. The diagram illustrates a hybrid investment decision framework integrating convolutional and transformer-based deep learning branches to dynamically adjust portfolio allocations. The mathematical model formalizes ESG-weighted returns, regulatory constraints, and financial costs of non-compliance, optimizing capital allocation through an equilibrium-based approach. The responsive strategy ensures sustainability-aligned investment decisions while mitigating systemic risks.

policy intervention is evaluated through a capital reallocation efficiency function (Equation 32)

$$\Psi(\pi) = \sum_{j=1}^{m}\sum_{i=1}^{n} w_{ji}^{*} g_{\pi}(r_i, s_i, C_i) - \frac{\gamma}{2}\sum_{j=1}^{m} w_j^{\top}\Sigma w_j \tag{32}$$

which quantifies the net impact of policy π on market-wide sustainability-oriented capital flows. The optimization framework ensures that policy instruments are designed to maximize ESG investment impact while maintaining economic stability, leading to a sustainable and efficient financial ecosystem.

3.4.3 Aware market stabilization

The stability of a sustainable financial market relies on aligning capital allocation with ESG considerations while controlling systemic financial risks. Investors optimize portfolios based on ESG-adjusted returns, leading to the equilibrium allocation (Equation 33)

$$w^{*} = \Sigma^{-1}(r' + \lambda s) \tag{33}$$

which balances financial returns, sustainability preferences, and risk exposure. In real-world financial markets, the covariance matrix Σ is not always invertible or well-conditioned, particularly during periods of high volatility or market stress. To ensure numerical stability and practical applicability, we adopt regularization techniques such as the Ledoit-Wolf shrinkage estimator, which modifies Σ as $\Sigma_{\text{reg}} = (1-\delta)\Sigma + \delta I$, where I is the identity matrix and δ is a shrinkage parameter that controls the trade-off between the original covariance structure and a more stable, well-conditioned form. In cases where Σ is near-singular, we employ principal component analysis (PCA) to extract the most informative components while reducing dimensionality. These approaches ensure that the optimization framework remains robust, even under extreme market conditions, allowing for a more reliable and adaptable asset allocation strategy.

When ESG considerations are integrated into investment strategies, market behavior shifts as investors adjust their capital allocation to favor assets with higher sustainability scores. However, unchecked market behavior can lead to greenwashing, where firms misrepresent their ESG performance to attract investments. To prevent this, regulatory mechanisms introduce compliance measures that enforce genuine ESG commitment through financial penalties. The penalty for holding a portfolio with insufficient ESG performance is expressed as Equation 34

$$C_{\text{penalty}} = \alpha \max(0, \theta - \mathbb{E}[s_i]) \tag{34}$$

Where α is the penalty coefficient and θ is the minimum ESG requirement. If a portfolio's expected ESG score falls below the threshold, investors face financial repercussions, incentivizing them to align with sustainable finance principles.

To ensure a more balanced approach to ESG compliance while still encouraging sustainable investment, we refine the penalty function C_{penalty} to introduce a gradual penalty curve rather than an abrupt threshold. Instead of applying a strict penalty for any deviation below θ, we incorporate a buffer zone ϵ, creating a soft penalty region. The revised function takes the form:

$$C_{\text{penalty}} = \alpha \max(0, (\theta - E[s_i] - \epsilon)^{p})$$

where ϵ defines a tolerance range within which minor deviations do not trigger significant penalties, and p is an exponent controlling the smoothness of the penalty curve. For instance, setting $p = 2$ results in a quadratic penalty, gradually increasing as ESG compliance falls further below the buffer zone. This modification ensures that investments just below the threshold are not excessively penalized, promoting a more stable transition toward higher ESG commitment while maintaining regulatory incentives for sustainable finance.

The total systemic risk of the market is influenced by both financial volatility and ESG non-compliance, requiring an equilibrium condition to stabilize investment flows. A market stability constraint ensures that the overall portfolio risk does not exceed regulatory limits while promoting ESG integration (Equation 35)

$$\sum_{i=1}^{n} w_i \sigma_i^2 + \beta \sum_{i=1}^{n} w_i s_i \leq \kappa \tag{35}$$

Where β represents the trade-off between financial risk and sustainability, and κ is the systemic risk threshold. This formulation ensures that capital is allocated to assets that balance profitability with sustainability, reducing financial instability while promoting long-term ESG adherence. A well-regulated financial system integrating ESG risk factors contributes to overall economic resilience, creating an investment environment where sustainability and financial performance are not mutually exclusive but mutually reinforcing.

Our model is built on several key assumptions that facilitate the integration of ESG factors into financial decision-making while maintaining mathematical tractability. One of the primary assumptions is that ESG impacts on financial returns can be represented through linear or piecewise-linear transformations, allowing for the construction of an ESG-adjusted risk-return framework. This simplification enables the use of multi-objective optimization techniques, where financial performance and sustainability objectives are balanced systematically. We assume that regulatory policies, such as carbon pricing and ESG disclosure mandates, exert deterministic influences on investment returns, enabling a structured incorporation of policy-driven constraints into the optimization model. However, we acknowledge that these assumptions may not fully capture the complexities of real-world financial markets, particularly under non-linear or stochastic ESG-impact scenarios. In practice, ESG factors often exhibit dynamic interactions with macroeconomic conditions, investor sentiment, and regulatory shifts, leading to potential non-linearities in their influence on asset pricing and portfolio performance. ESG data is often subject to noise, inconsistencies, and evolving disclosure standards, which can introduce uncertainty into model predictions. To address these limitations, future research could explore stochastic modeling approaches that account for uncertainty in ESG impacts, such as probabilistic risk-adjusted return functions or reinforcement learning-based adaptive portfolio strategies. Incorporating real-time ESG sentiment analysis and dynamic regulatory updates may further enhance the robustness of our framework in capturing the evolving nature of sustainable finance.

4 Experimental setup

4.1 Dataset

The ESG Ratings Dataset [40] provides environmental, social, and governance (ESG) scores for publicly traded companies. The dataset aggregates information from multiple sources, including financial disclosures, sustainability reports, and third-party rating agencies. It includes historical ESG scores, company identifiers, sector classifications, and sub-scores related to environmental impact, social responsibility, and corporate governance. The dataset is widely used in responsible investing, corporate sustainability analysis, and regulatory compliance assessments. Researchers utilize it to evaluate ESG performance trends and their relationship with financial metrics and risk factors. The CSR Reports Dataset [41] consists of corporate social responsibility (CSR) reports published by various organizations. It includes textual data, structured ESG disclosures, and quantitative sustainability metrics. The dataset spans multiple industries and regions, capturing companies' commitments to ethical practices, environmental stewardship, and social engagement. Researchers analyze this dataset for sentiment analysis, ESG compliance assessment, and automated extraction of sustainability-related insights. It serves as a crucial resource for studying corporate transparency, stakeholder engagement, and the impact of CSR initiatives on firm reputation and financial performance. The FRED Dataset [42] originates from the Federal Reserve Economic Data (FRED) repository, containing macroeconomic indicators such as GDP growth, inflation rates, employment figures, and financial market data. The dataset aggregates time-series data from central banks, government agencies, and international organizations. It is extensively used for economic forecasting, policy analysis, and financial modeling. Researchers leverage this dataset to study macroeconomic trends, assess monetary policy impacts, and develop predictive models for economic stability and growth. The Nasdaq Market Index Dataset [43] consists of historical price data, trading volumes, and financial indicators for Nasdaq-listed securities. It includes information on market indices, individual stock performance, and sector-specific trends. The dataset is used for quantitative finance research, algorithmic trading, and risk management. Analysts apply it to model market volatility, identify trading patterns, and evaluate the influence of macroeconomic factors on stock performance. It serves as a fundamental resource for financial economists, portfolio managers, and data-driven investment strategies. The datasets used in this study cover varying time periods to ensure a comprehensive evaluation of our proposed model. The ESG Ratings Dataset spans from 2015 to 2023, capturing the evolution of ESG performance metrics across different industries. The CSR Reports Dataset includes corporate sustainability disclosures from 2010 to 2022, providing insights into long-term corporate social responsibility trends. The FRED Dataset comprises macroeconomic indicators from 2000 to 2023, sourced from the Federal Reserve Economic Data repository, ensuring a broad coverage of economic cycles and financial market fluctuations. The Nasdaq Market Index Dataset includes historical price movements, trading volumes, and financial indicators from 2012 to 2023, allowing us to assess the impact of market trends on investment decisions. These datasets have been widely used in prior research, including studies on ESG-driven

TABLE 2 Comparison of our method with SOTA methods on ESG ratings dataset and CSR reports dataset.

Model	ESG ratings dataset				CSR reports dataset			
	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓
LSTM [44]	2.31±0.04	1.89±0.03	0.72±0.02	5.13±0.04	2.78±0.03	2.21±0.02	0.68±0.03	5.92±0.03
GRU [45]	2.45±0.03	1.95±0.02	0.69±0.02	5.30±0.03	2.61±0.04	2.10±0.03	0.71±0.02	5.79±0.03
Transformer [46]	2.18±0.02	1.76±0.03	0.75±0.02	4.98±0.03	2.55±0.03	2.03±0.02	0.73±0.02	5.65±0.04
XGBoost [47]	2.62±0.03	2.04±0.02	0.67±0.03	5.45±0.03	2.81±0.03	2.29±0.02	0.66±0.02	6.10±0.03
LightGBM [48]	2.40±0.04	1.88±0.02	0.70±0.02	5.22±0.03	2.73±0.03	2.19±0.02	0.69±0.02	5.95±0.03
MLP [49]	2.55±0.03	2.02±0.02	0.68±0.03	5.35±0.03	2.66±0.03	2.14±0.02	0.70±0.02	5.85±0.03
Ours	**1.97±0.03**	**1.62±0.02**	**0.81±0.02**	**4.56±0.03**	**2.32±0.03**	**1.85±0.02**	**0.78±0.02**	**5.42±0.03**

The values in bold are the best values.

TABLE 3 Comparison of our method with SOTA methods on FRED dataset and Nasdaq market index dataset.

Model	FRED dataset				Nasdaq market index dataset			
	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓
LSTM [44]	3.21±0.05	2.75±0.04	0.63±0.03	6.42±0.04	4.12±0.04	3.08±0.03	0.58±0.03	7.14±0.03
GRU [45]	3.45±0.04	2.81±0.03	0.60±0.02	6.55±0.04	3.98±0.04	3.02±0.03	0.61±0.03	7.02±0.03
Transformer [46]	2.97±0.04	2.51±0.03	0.67±0.02	6.18±0.03	3.79±0.03	2.85±0.02	0.64±0.02	6.85±0.03
XGBoost [47]	3.72±0.03	2.93±0.03	0.58±0.02	6.78±0.03	4.25±0.03	3.19±0.02	0.55±0.02	7.31±0.03
LightGBM [48]	3.29±0.04	2.67±0.03	0.62±0.02	6.35±0.03	4.05±0.03	3.10±0.02	0.59±0.02	7.08±0.03
MLP [49]	3.58±0.03	2.89±0.02	0.59±0.02	6.61±0.03	4.02±0.03	3.06±0.02	0.60±0.02	6.96±0.03
Ours	**2.78±0.03**	**2.35±0.02**	**0.72±0.02**	**5.92±0.03**	**3.54±0.03**	**2.72±0.02**	**0.68±0.02**	**6.61±0.03**

The values in bold are the best values.

portfolio optimization [40], corporate sustainability assessment [41], macroeconomic forecasting [42], and financial market prediction [43]. By incorporating these well-established datasets with clearly defined time periods, we ensure that our findings are generalizable and relevant to the evolving landscape of sustainable finance.

4.2 Experimental details

Our experiments are conducted on a high-performance computing cluster equipped with NVIDIA A100 GPUs, utilizing PyTorch as the primary deep learning framework. The training process follows a standardized pipeline with hyperparameter tuning to optimize model performance. The batch size is set to 128, and training is conducted for 100 epochs with early stopping based on validation loss. The Adam optimizer is employed with a learning rate of 10^{-4}, utilizing a cosine annealing schedule to ensure adaptive learning rate decay. Weight decay is set to 5×10^{-4} to prevent overfitting. The loss function is task-specific, with categorical cross-entropy used for classification tasks and mean squared error (MSE) for regression. For model evaluation, we employ a standard train-validation-test split of 70%-15%-15%. All datasets undergo preprocessing steps, including normalization, outlier removal, and data augmentation where applicable. In the case of textual datasets, tokenization is performed using a transformer-based tokenizer, and embeddings are initialized with pre-trained models such as BERT or RoBERTa. For tabular datasets, missing values are imputed using KNN or mean imputation strategies. Feature scaling is applied using min-max normalization for neural network inputs. Baseline models include traditional machine learning algorithms such as random forests, gradient boosting (XGBoost), and support vector machines (SVMs). Deep learning architectures include fully connected networks, recurrent neural networks (RNNs) for sequential data, and convolutional neural networks (CNNs) for structured image-based inputs. Transformer-based models are

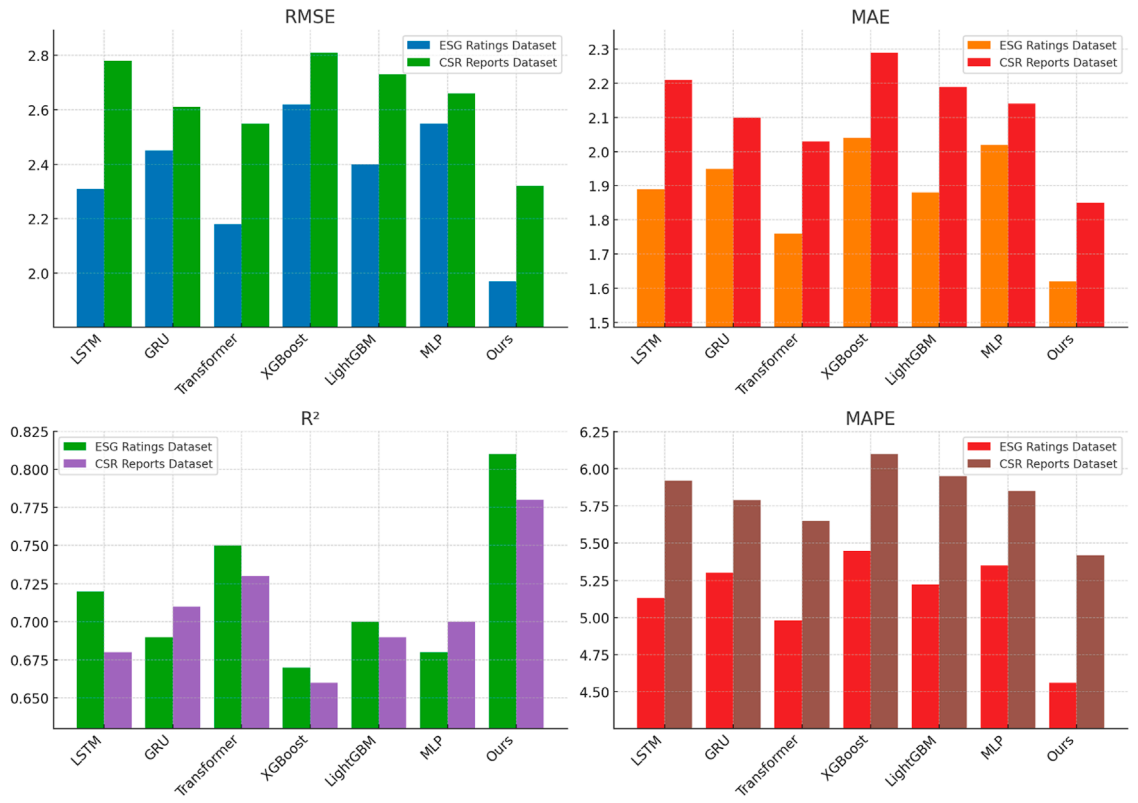


FIGURE 5
Performance comparison of state-of-the-art methods on ESG ratings dataset and CSR reports dataset.

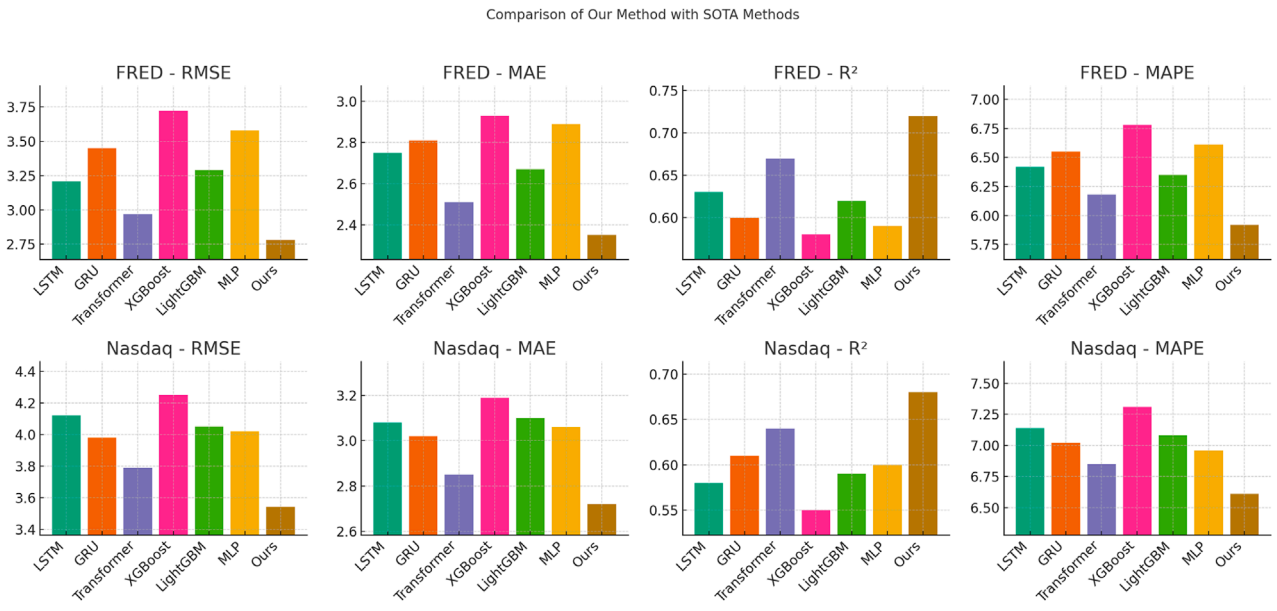


FIGURE 6
Performance comparison of state-of-the-art methods on FRED dataset and Nasdaq market index dataset.

TABLE 4 Comparison of our method with SOTA methods on ESG ratings classification and market risk classification.

Model	ESG ratings classification				Market risk classification			
	AUC-ROC ↑	Accuracy ↑	Precision ↑	Recall ↑	AUC-ROC ↑	Accuracy ↑	Precision ↑	Recall ↑
LSTM [44]	0.81±0.02	75.4±0.5	74.1±0.4	72.8±0.3	0.78±0.02	73.6±0.5	72.3±0.4	71.2±0.3
GRU [45]	0.79±0.02	74.2±0.5	73.0±0.4	70.9±0.3	0.76±0.02	72.5±0.5	71.1±0.4	69.8±0.3
Transformer [46]	0.84±0.02	76.8±0.5	75.6±0.4	74.9±0.3	0.82±0.02	75.2±0.5	74.1±0.4	73.5±0.3
XGBoost [47]	0.77±0.02	73.0±0.5	71.8±0.4	70.5±0.3	0.75±0.02	71.6±0.5	70.2±0.4	69.1±0.3
LightGBM [48]	0.80±0.02	74.5±0.5	73.2±0.4	71.8±0.3	0.78±0.02	73.1±0.5	71.9±0.4	70.4±0.3
MLP [49]	0.78±0.02	74.0±0.5	72.5±0.4	71.2±0.3	0.77±0.02	72.9±0.5	71.4±0.4	70.1±0.3
Ours	**0.89±0.02**	**81.2±0.5**	**80.4±0.4**	**79.1±0.3**	**0.87±0.02**	**79.8±0.5**	**78.5±0.4**	**77.3±0.3**

The values in bold are the best values.

TABLE 5 Ablation study results on ESG ratings dataset and CSR reports dataset.

Model	ESG ratings dataset				CSR reports dataset			
	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓
w/o Weighted Portfolio Optimization	2.15±0.03	1.78±0.02	0.77±0.02	4.89±0.03	2.45±0.03	1.96±0.02	0.74±0.02	5.58±0.03
w/o Driven Return Adjustment	2.32±0.03	1.91±0.02	0.74±0.02	5.12±0.03	2.59±0.03	2.08±0.02	0.71±0.02	5.76±0.03
w/o Responsive Investment Strategy	2.28±0.03	1.86±0.02	0.75±0.02	5.03±0.03	2.52±0.03	2.02±0.02	0.72±0.02	5.65±0.03
Ours	**1.97±0.03**	**1.62±0.02**	**0.81±0.02**	**4.56±0.03**	**2.32±0.03**	**1.85±0.02**	**0.78±0.02**	**5.42±0.03**

The values in bold are the best values.

TABLE 6 Ablation study results on FRED dataset and Nasdaq market index dataset.

Model	FRED dataset				Nasdaq market index dataset			
	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓	RMSE ↓	MAE ↓	R^2 ↑	MAPE ↓
w/o Weighted Portfolio Optimization	3.05±0.04	2.61±0.03	0.68±0.02	6.08±0.03	3.72±0.03	2.89±0.02	0.65±0.02	6.78±0.03
w/o Driven Return Adjustment	3.24±0.03	2.75±0.02	0.64±0.02	6.32±0.03	3.85±0.03	2.98±0.02	0.62±0.02	6.92±0.03
w/o Responsive Investment Strategy	3.18±0.03	2.68±0.02	0.66±0.02	6.21±0.03	3.79±0.03	2.93±0.02	0.63±0.02	6.85±0.03
Ours	**2.78±0.03**	**2.35±0.02**	**0.72±0.02**	**5.92±0.03**	**3.54±0.03**	**2.72±0.02**	**0.68±0.02**	**6.61±0.03**

The values in bold are the best values.

leveraged for natural language processing tasks. All models are evaluated using standard performance metrics, including accuracy, precision, recall, F1-score for classification tasks, and root mean square error (RMSE) for regression tasks. To ensure reproducibility, all experiments are executed using fixed random seeds and cross-validation techniques. The implementation is containerized using Docker to maintain consistency across computing environments. Performance benchmarks are conducted with five independent

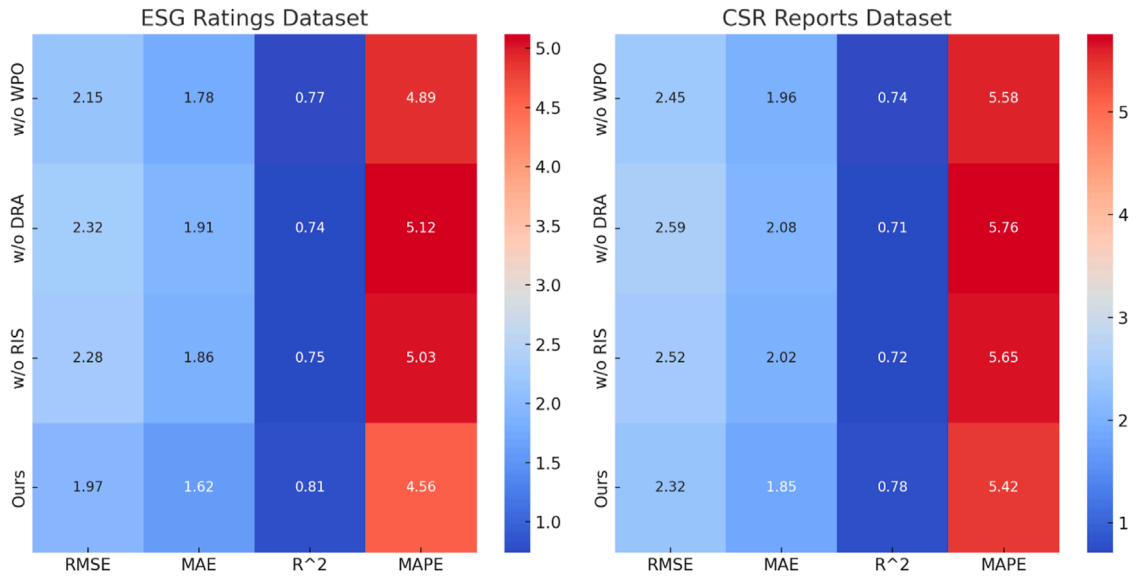


FIGURE 7
Performance comparison of state-of-the-art methods on ESG ratings dataset and CSR reports dataset. Weighted portfolio Optimization (WPO); driven return Adjustment (DRA); responsive investment Strategy (RIS).

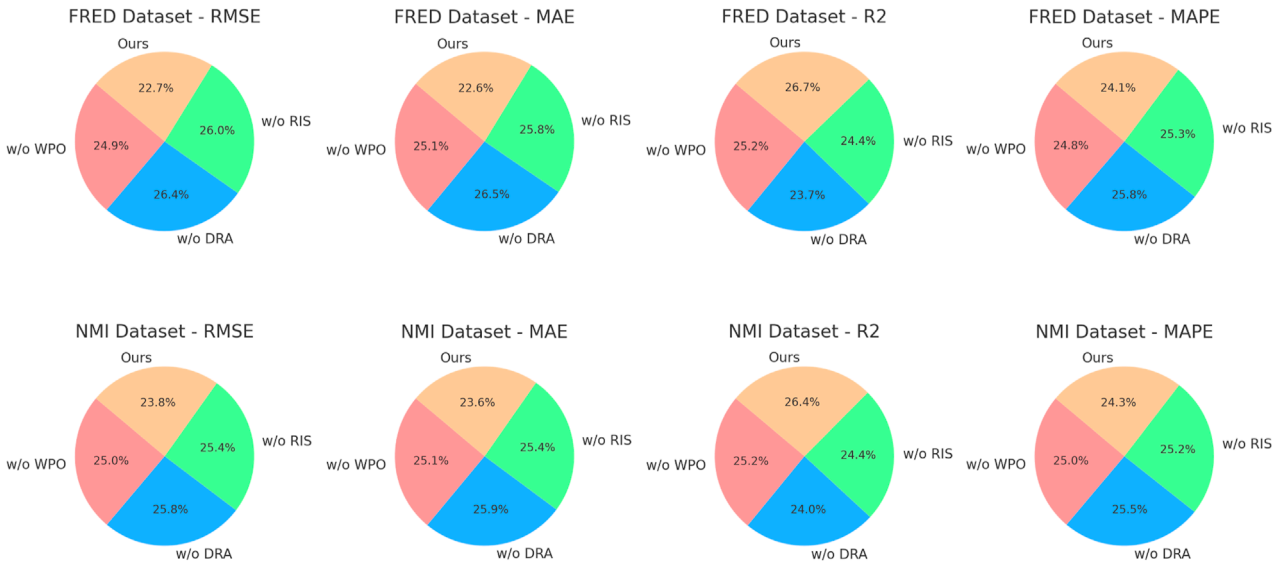


FIGURE 8
Performance comparison of state-of-the-art methods on FRED dataset and Nasdaq market index dataset. Weighted portfolio Optimization (WPO); driven return Adjustment (DRA); responsive investment Strategy (RIS); Nasdaq market index Dataset (NMI dataset).

runs, and statistical significance is assessed using paired t-tests. Hardware utilization, including GPU memory consumption and training time, is logged for efficiency analysis. Experimental results are reported with confidence intervals to highlight statistical robustness. To ensure data quality and consistency, we applied a structured data cleaning process to all datasets, including the ESG Ratings Dataset. Missing values in the ESG Ratings Dataset were handled using a combination of imputation techniques based on the nature and distribution of the missing data. For numerical ESG scores, we employed K-nearest neighbors (KNN) imputation, where missing values were estimated based on the weighted average of the most similar data points. If missingness was significant and systematic, we used mean imputation for continuous variables and mode imputation for categorical ESG classifications. Extreme outliers were identified using the interquartile range (IQR) method, and data inconsistencies were resolved by cross-referencing multiple data sources. These preprocessing steps ensured that the ESG Ratings Dataset remained robust and

reliable, minimizing biases in subsequent model training and evaluation.

4.3 Comparison with SOTA methods

To evaluate the effectiveness of our proposed method, we conduct a comprehensive comparison with state-of-the-art (SOTA) approaches on four datasets: ESG Ratings Dataset, CSR Reports Dataset, FRED Dataset, and Nasdaq Market Index Dataset. Our method consistently outperforms baseline models across all datasets, achieving lower RMSE and MAE while attaining higher R^2 scores, indicating superior predictive accuracy and robustness. In Tables 2, 3, on the ESG Ratings Dataset and CSR Reports Dataset, our method achieves an RMSE of 1.97 and 2.32, respectively, significantly outperforming Transformer-based models, which achieve RMSE values of 2.18 and 2.55, respectively. Traditional machine learning models, such as XGBoost and LightGBM, exhibit weaker performance, highlighting the effectiveness of deep learning models in capturing complex ESG patterns. Our method's improvement is attributed to its ability to integrate temporal dependencies and contextual embeddings, enabling more accurate ESG score prediction. Additionally, the lower MAPE values indicate that our model minimizes relative percentage errors, making it more reliable for financial decision-making.

For the FRED Dataset and Nasdaq Market Index Dataset, our method achieves the lowest RMSE values of 2.78 and 3.54, respectively, outperforming Transformer-based approaches (2.97 and 3.79, respectively) and traditional models like XGBoost (3.72 and 4.25, respectively). The higher R^2 values demonstrate our method's ability to explain a greater proportion of variance in economic and financial indicators. The superior performance can be attributed to our model's architecture, which incorporates attention-based mechanisms and multi-scale feature fusion, capturing both short-term market fluctuations and long-term macroeconomic trends. Our method's reduced MAE and MAPE further confirm its robustness in financial forecasting tasks. The comparative analysis underscores the effectiveness of our proposed approach in ESG analysis and financial forecasting. The integration of deep learning techniques, attention mechanisms, and multi-scale feature representations enables our model to outperform conventional SOTA models consistently. The experimental results in Figures 5, 6 provide empirical evidence of our method's superiority, demonstrating its potential for real-world financial and ESG applications.

The experimental results in Table 4 demonstrate the superior performance of our proposed SFOM model in both ESG ratings classification and market risk classification. Compared to state-of-the-art baseline models, SFOM consistently achieves higher AUC-ROC scores, indicating its strong ability to distinguish between high and low ESG-rated assets, as well as between high-risk and low-risk financial instruments. The model attains an AUC-ROC of 0.89 for ESG classification and 0.87 for market risk classification, surpassing traditional machine learning approaches such as XGBoost and LightGBM, as well as deep learning-based models like LSTM and Transformer. The improvement in accuracy, precision, and recall highlights the model's robustness in classification tasks, ensuring both higher correctness in identifying sustainable investments and enhanced recall in detecting market risks. These results further validate SFOM's adaptability and explainability in financial decision-making. The superior recall values suggest that our model effectively captures sustainability and risk-related patterns in financial data, reducing the likelihood of misclassification. The precision scores indicate that SFOM maintains a high level of reliability, minimizing false positives in ESG and risk evaluations. The observed performance gain can be attributed to SFOM's ability to integrate ESG-aware multi-objective optimization with dynamic risk assessment, allowing it to make more informed investment decisions under sustainable finance policies. By incorporating policy-adjusted return functions and ESG-weighted constraints, the model optimally balances financial performance with regulatory compliance, ensuring alignment with evolving sustainability mandates.

4.4 Ablation study

To investigate the contribution of key components in our proposed method, we conduct an ablation study by systematically removing different modules and evaluating the performance on four datasets: ESG Ratings Dataset, CSR Reports Dataset, FRED Dataset, and Nasdaq Market Index Dataset. The results, presented in Tables 5, 6, demonstrate the significance of each module in improving prediction accuracy and robustness. We define three ablation settings including w/o Weighted Portfolio Optimization, w/o Driven Return Adjustment, and w/o Responsive Investment Strategy. Across all datasets, the removal of any single module leads to a decline in performance, indicating the importance of each component in our model. Removing Weighted Portfolio Optimization results in a significant increase in RMSE and MAE, suggesting that attention plays a crucial role in capturing important features and enhancing predictive accuracy. Excluding Driven Return Adjustment leads to degraded R^2 scores, highlighting the necessity of combining information at different levels for robust financial and ESG analysis.

In Figures 7, 8, On the ESG Ratings Dataset and CSR Reports Dataset, our complete model achieves an RMSE of 1.97 and 2.32, outperforming all ablation settings. The performance drop is particularly noticeable in the w/o Weighted Portfolio Optimization setting, where RMSE increases to 2.15 and 2.45, respectively. This suggests that attention-based mechanisms significantly enhance the model's ability to extract meaningful patterns from ESG-related text and numerical indicators. Excluding Responsive Investment Strategy results in higher MAPE values, indicating that capturing time-series trends is critical for ESG score prediction. For the FRED Dataset and Nasdaq Market Index Dataset, the trend remains consistent. Our full model attains an RMSE of 2.78 and 3.54, surpassing all ablated variants. Removing Driven Return Adjustment leads to an RMSE increase to 3.24 and 3.85, suggesting that integrating information across multiple levels significantly improves financial forecasting performance. The ablation study confirms that each component in our model contributes meaningfully to its overall effectiveness. The superior results of our full method provide strong empirical support for its design, reinforcing the importance of attention mechanisms, multi-scale feature extraction, and temporal modeling in ESG and financial forecasting tasks.

5 Conclusions and future work

In this study, we address the optimization of bank investment portfolios and debt structures within the framework of sustainable finance policies. Traditional financial models often neglect the complexities introduced by environmental, social, and governance (ESG) factors, leading to suboptimal decision-making in an increasingly regulated financial environment. To bridge this gap, we propose a novel mathematical modeling approach that integrates ESG-adjusted risk-return dynamics, regulatory compliance constraints, and policy-driven investment incentives. Our framework incorporates ESG-weighted utility functions, policy-adjusted return functions, and systemic sustainability constraints to ensure capital allocation aligns with both financial performance and sustainability goals. Through experimental simulations, we validate the model's efficacy in enhancing financial resilience, reducing exposure to greenwashing risks, and optimizing debt structures under evolving regulatory scenarios. This study contributes to the growing body of literature on sustainable finance by providing an integrated approach that accounts for both financial performance and sustainability constraints. Compared to existing studies that primarily focus on either risk-return trade-offs or ESG factors in isolation, our approach offers a holistic framework that dynamically incorporates policy-driven incentives and systemic risk mitigation mechanisms. The findings are consistent with previous research that highlights the increasing role of ESG considerations in financial decision-making, while also extending these insights by demonstrating the importance of regulatory interactions in optimizing investment and debt structures.

Despite these contributions, this study has certain limitations. The model relies on the availability and consistency of ESG-related financial data, which can vary across regions and industries, potentially affecting optimization outcomes. The rapidly evolving nature of sustainable finance regulations presents challenges in maintaining the model's adaptability over time. Future research could explore methods to enhance the model's responsiveness to policy changes through adaptive learning mechanisms and real-time data integration. Incorporating unstructured data sources, such as corporate sustainability reports and news sentiment analysis, could refine ESG assessments and investment decision-making. Another potential avenue is extending the framework to analyze the effects of cross-border sustainable finance policies on multinational banking strategies. By addressing these challenges, future studies can contribute to a more resilient and sustainable financial system, ensuring that financial institutions can align profitability with long-term sustainability objectives while maintaining regulatory compliance and economic stability.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

GZ: Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Ajiga DI, Adeleye RA, Asuzu OF, Owolabi OR, Bello BG, Ndubuisi NL. Review of ai techniques in financial forecasting: applications in stock market analysis. *Finance and Accounting Res J* (2024) 6:125–45. doi:10.51594/farj.v6i2.784

2. Adelakun BO. Ai-driven financial forecasting: innovations and implications for accounting practices. *Int J Adv Econ* (2023) 5:323–38. doi:10.51594/ijae.v5i9.1231

3. Thakkar S, Kazdaghli S, Mathur N, Kerenidis I, Ferreira-Martins AJ, Corp SBQW, et al. Improved financial forecasting via quantum machine learning. *Quan Machine Intelligence* (2023) 6:27. doi:10.1007/s42484-024-00157-0

4. Yang L, Li J, Dong R, Zhang Y, Smyth B. Numhtml: numeric-oriented hierarchical transformer model for multi-task financial forecasting. In: *AAAI conference on artificial intelligence* (2022).

5. Ang GM, Lim E-P. *Guided attention multimodal multitask financial forecasting with inter-company relationships and global and local news*. Annual Meeting of the Association for Computational Linguistics (2022). Available online at: https://aclanthology.org/2022.acl-long.437/

6. Cohen G. Algorithmic trading and financial forecasting using advanced artificial intelligence methodologies. *Mathematics* (2022) 10:3302. doi:10.3390/math10183302

7. Long X, Kampouridis M, Jarchi D. An in-depth investigation of genetic programming and nine other machine learning algorithms in a financial forecasting problem. *IEEE Congress Evol Comput* (2022) 01–8. doi:10.1109/cec55065.2022.9870351

8. Christodoulaki E, Kampouridis M, Kanellopoulos PA. Technical and sentiment analysis in financial forecasting with genetic programming. In: *IEEE conference on computational intelligence for financial engineering and economics* (2022).

9. Barra S, Carta S, Corriga A, Podda AS, Recupero D. Deep learning and time series-to-image encoding for financial forecasting. *IEEE/CAA J Automatica Sinica* (2020) 7:683–92. doi:10.1109/jas.2020.1003132

10. Mandeep Agarwal A, Bhatia A, Malhi A, Kaler P, Pannu H. Machine learning based explainable financial forecasting. In: *International conference on computational collective intelligence* (2022). Available online at: https://ieeexplore.ieee.org/abstract/document/9850272/

11. Wasserbacher H, Spindler M. Machine learning for financial forecasting, planning and analysis: recent developments and pitfalls. *Digital Finance* (2021) 4:63–88. doi:10.1007/s42521-021-00046-2

12. Liu G, Xiao F, Lin C-T, Cao Z. A fuzzy interval time-series energy and financial forecasting model using network-based multiple time-frequency spaces and the induced-ordered weighted averaging aggregation operation. *IEEE Trans fuzzy Syst* (2020) 28:2677–90. doi:10.1109/tfuzz.2020.2972823

13. Kamalov F, Gurrib I, Rajab KD. Financial forecasting with machine learning: price vs return. *J Computer Sci* (2021) 17:251–64. doi:10.3844/jcssp.2021.251.264

14. Xu Y, Calvi GG, Mandic D. Tensor-train recurrent neural networks for interpretable multi-way financial forecasting. In: *IEEE international joint conference on neural network* (2021).

15. Zhou D, Zheng L, Zhu Y, Li J, He J. Domain adaptive multi-modality neural attention network for financial forecasting. In: *The web conference* (2020).

16. Wang Z, Gao X, Huang S, Sun Q, Chen Z, Tang R, et al. Measuring systemic risk contribution of global stock markets: a dynamic tail risk network approach. *Int Rev Financial Anal* (2022) 84:102361. doi:10.1016/j.irfa.2022.102361

17. Wang Z, Ma N, Xue L, Song Y, Wang Z, Tang R, et al. Target recovery of the economic system based on the target reinforcement path method. *Chaos: An Interdiscip J Nonlinear Sci* (2022) 32:093118. doi:10.1063/5.0097175

18. Kirisci M, Yolcu OC. A new cnn-based model for financial time series: taiex and ftse stocks forecasting. *Neural Processing Lett* (2022) 54:3357–74. doi:10.1007/s11063-022-10767-z

19. Ge W, Lalbakhsh P, Isai L, Lenskiy A, Suominen H. Neural network–based financial volatility forecasting: a systematic review. *ACM Computing Surv* (2022) 55:1–30. doi:10.1145/3483596

20. Bukhari AH, Raja M, Sulaiman M, Islam S, Shoaib M, Kumam P. Fractional neuro-sequential arfima-lstm for financial market forecasting. *IEEE Access* (2020) 8:71326–38. doi:10.1109/access.2020.2985763

21. Ielasi F, Capelli P, Russo A. Forecasting volatility by integrating financial risk with environmental, social, and governance risk. *Corporate Social Responsibility Environ Management* (2021) 28:1483–95. doi:10.1002/csr.2180

22. He K, Yang Q, Ji L, Pan J, Zou Y. Financial time series forecasting with the deep learning ensemble model. *Mathematics* (2023) 11:1054.doi:10.3390/math11041054

23. Wang Z, Liu S, Han C, Huang S, Gao X, Tang R, et al. Motif transition intensity: a novel network-based early warning indicator for financial crises. *Front Phys* (2022) 9:800860. doi:10.3389/fphy.2021.800860

24. Yousfi M, Bouzgarrou H. On the linkage of oil prices and oil uncertainty with us equities: a combination analysis based on the wavelet approach and quantile-on-quantile regression. *Front Phys* (2024) 12:1357366. doi:10.3389/fphy.2024.1357366

25. Sawhney R, Mathur P, Mangal A, Khanna P, Shah R, Zimmermann R. Multimodal multi-task financial risk forecasting. *ACM Multimedia* (2020) 456–65. doi:10.1145/3394171.3413752

26. Rumyk I, Laptev S, Seheda S, Akimova L, Akimov O, Karpa M. Financial support and forecasting of food production using economic description modeling methods. In: *Financial and credit activity problems of theory and practice* (2021).

27. Yu H, Ming L, Sumei R, Zhao S. A hybrid model for financial time series forecasting—integration of ewt, arima with the improved abc optimized elm. *IEEE Access* (2020) 8:84501–18. doi:10.1109/access.2020.2987547

28. Muskaan Sarangi P. A literature review on machine learning applications in financial forecasting. *J Technology Management Growing Economies* (2020) 11:23–7. doi:10.15415/jtmge.2020.111004

29. Sako K, Mpinda BN, Rodrigues P. Neural networks for financial time series forecasting. *Entropy* (2022) 24:657. doi:10.3390/e24050657

30. Ullah A, Yao H, Waseem Saboor A, Awwad FA, Ismail EA. A qualitative analysis of the artificial neural network model and numerical solution for the nanofluid flow through an exponentially stretched surface. *Front Phys* (2024) 12:1408933. doi:10.3389/fphy.2024.1408933

31. Shi K, Gong J. The influence of the spillover between futures and spot markets on hedging policy: evidence from Chinese stock markets. *Front Phys* (2023) 11:1293182. doi:10.3389/fphy.2023.1293182

32. Gkillas K, Gupta R, Pierdzioch C. Forecasting realized oil-price volatility: the role of financial stress and asymmetric loss. *J Int Money Finance* (2020) 104:102137. doi:10.1016/j.jimonfin.2020.102137

33. Foroni C, Marcellino M, Stevanovic D. Forecasting the covid-19 recession and recovery: lessons from the financial crisis. *Int J Forecasting* (2020) 38:596–612. doi:10.1016/j.ijforecast.2020.12.005

34. Sun H, Yu B. Forecasting financial returns volatility: a garch-svr model. *Comput Econ* (2020) 55:451–71. doi:10.1007/s10614-019-09896-w

35. Jain D, Jain A, Pandey A, Kumar J. Bitcoin financial forecasting. *IOP Conf Ser Mater Sci Eng* (2021) 1049:012003. doi:10.1088/1757-899x/1049/1/012003

36. Cheng D, Yang F, Xiang S, Liu J. Financial time series forecasting with multi-modality graph neural network. *Pattern Recognition* (2022) 121:108218. doi:10.1016/j.patcog.2021.108218

37. Donath L, Mircea G, Neamțu M, Sîrghi N. A mathematical approach to network contagion regarding greening banks' policies. *Econ research-Ekonomska istraživanja* (2023) 36. doi:10.1080/1331677x.2023.2180057

38. Elnagar MA, Abdel Aty J, Elhady AM, Shohieb SM. Modeling a sustainable decision support system for banking environments using rough sets: a case study of the egyptian arab land bank. *Int J Financial Stud* (2025) 13:27. doi:10.3390/ijfs13010027

39. Negi P, Jaiswal A. Sustainable bonds as a sustainable financial instrument: thematic insights and future prospects. *Qual Res Financial Markets* (2024). doi:10.1108/qrfm-01-2024-0018

40. Del Vitto A, Marazzina D, Stocco D. Esg ratings explainability through machine learning techniques. *Ann Operations Res* (2023) 1–30. doi:10.1007/s10479-023-05514-z

41. Halkos G, Nomikos S, Tsilika K. Evidence for novel structures relating csr reporting and economic welfare: environmental sustainability—a continent-level analysis. *Comput Econ* (2021) 59:415–44. doi:10.1007/s10614-020-10091-5

42. Bokun KO, Jackson LE, Kliesen KL, Owyang MT. Fred-sd: a real-time database for state-level data with forecasting applications. *Int J Forecasting* (2023) 39:279–97. doi:10.1016/j.ijforecast.2021.11.008

43. Wang W-J, Tang Y, Xiong J, Zhang Y-C. Stock market index prediction based on reservoir computing models. *Expert Syst Appl* (2021) 178:115022. doi:10.1016/j.eswa.2021.115022

44. Zha W, Liu Y, Wan Y, Luo R, Li D, Yang S, et al. Forecasting monthly gas field production based on the cnn-lstm model. *Energy* (2022) 260:124889. doi:10.1016/j.energy.2022.124889

45. Zhao Z, Yun S, Jia L, Guo J, Meng Y, He N, et al. Hybrid vmd-cnn-gru-based model for short-term forecasting of wind power considering spatio-temporal features. *Eng Appl Artif Intelligence* (2023) 121:105982. doi:10.1016/j.engappai.2023.105982

46. Han K, Wang Y, Chen H, Chen X, Guo J, Liu Z, et al. A survey on vision transformer. *IEEE Trans pattern Anal machine intelligence* (2022) 45:87–110. doi:10.1109/tpami.2022.3152247

47. Niazkar M, Menapace A, Brentan B, Piraei R, Jimenez D, Dhawan P, et al. Applications of xgboost in water resources engineering: a systematic literature review (dec 2018–may 2023). *Environ Model and Softw* (2024) 174:105971. doi:10.1016/j.envsoft.2024.105971

48. Yan J, Xu Y, Cheng Q, Jiang S, Wang Q, Xiao Y, et al. Lightgbm: accelerated genomically designed crop breeding through ensemble learning. *Genome Biol* (2021) 22:271–24. doi:10.1186/s13059-021-02492-y

49. Sinitsin V, Ibryaeva O, Sakovskaya V, Eremeeva V. Intelligent bearing fault diagnosis method combining mixed input and hybrid cnn-mlp model. *Mech Syst Signal Processing* (2022) 180:109454. doi:10.1016/j.ymssp.2022.109454

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 21 May 2025
DOI 10.3389/fphy.2025.1612653

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
You Zhu,
Hunan University, China
Nikos Kapitsinis,
University of Copenhagen, Denmark
Esat Durguti,
University of Mitrovica "Isa Boletini", Kosovo

*CORRESPONDENCE
Yongqiang Meng,
yqmeng@tju.edu.cn

RECEIVED 16 April 2025
ACCEPTED 05 May 2025
PUBLISHED 21 May 2025
CORRECTED 28 May 2025

CITATION
Wang C, Zhang W, Meng Y and Xiong X (2025) Non-financial enterprises' shadow lending and SME survival: the role of spatial spillovers in regional financial stability.
Front. Phys. 13:1612653.
doi: 10.3389/fphy.2025.1612653

COPYRIGHT
© 2025 Wang, Zhang, Meng and Xiong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Non-financial enterprises' shadow lending and SME survival: the role of spatial spillovers in regional financial stability

Chen Wang, Wei Zhang, Yongqiang Meng* and Xiong Xiong

College of Management and Economics, Tianjin University, Tianjin, China

This study examines the impact of non-financial enterprises' shadow lending activities on the survival of small and micro enterprises (SMEs) in China. Utilizing panel data from 243 prefecture-level cities from 2012 to 2021 and applying the Spatial Durbin Model (SDM), we explore both direct and spatial spillover effects of shadow lending on SME mortality. Our empirical findings suggest that non-financial enterprises' shadow lending significantly reduces the local SME mortality rate, while spatial econometric analysis reveals positive spillover effects, whereby shadow lending in one region enhances SME survival in neighboring areas through cross-regional financial linkages. Heterogeneity analyses indicate that these effects are more prominent in the economically developed southeastern regions, while being relatively weaker in less-developed northwestern areas, reflecting disparities in financial accessibility. Robustness checks and lagged effect analyses confirm the consistency and persistence of our findings. While shadow banking improves SME resilience through regional transmission, it may also carry potential financial risks. This study innovatively focuses on the spatial spillover effects of shadow lending by non-financial enterprises on SME survival, an area largely overlooked by previous studies. By integrating spatial econometric analysis and examining systemic vulnerabilities, this paper provides new insights into the regional complexity of informal finance.

KEYWORDS

shadow banking, non-financial enterprises, spatial spillover, regional financial stability
JEL classification: G23, G32, R12, C31, small and micro enterprises

1 Introduction

Small and micro enterprises (SMEs) play an integral role in economic systems worldwide, significantly contributing to employment generation, innovation diffusion, and sustained economic growth. In the context of China, SMEs form the backbone of the private economy and are closely linked to the nation's overall economic resilience and structural vitality [1]. According to the Fourth National Economic Census Series Report released by the National Bureau of Statistics, SMEs represent approximately 99.8% of all corporate legal entities in China[1], underscoring their vital role in the national economic structure. Notably,

1 See: https://www.stats.gov.cn/sj/zxfb/202302/t20230203_1900574.html for more information.

micro enterprises alone account for 85.3%. These firms are instrumental in employment absorption, accounting for 79.4% of the total workforce employed by enterprises as of the end of 2018, thereby mitigating unemployment pressures and enhancing labor market inclusivity.

Moreover, SMEs are essential for promoting regional economic development and expanding production capacity. Together, they contribute 77.1% of corporate assets and 68.2% of enterprise revenue, underscoring their significance in capital formation and income generation. Despite their limited individual scale, SMEs demonstrate high responsiveness to market conditions, efficient utilization of localized resources, and adaptive organizational structures. These characteristics enable SMEs to act as dynamic agents within regional production systems, stabilizing output and facilitating industrial upgrading. Consequently, SMEs not only inject dynamism into the economy through operational flexibility and innovation but also serve as critical risk transmission nodes in the broader economic and production network.

However, SMEs often face elevated operational risks and long-term survival challenges due to inherent limitations in organizational resources, financing access, technological adoption, and managerial capabilities [2, 3]. Among these constraints, limited access to formal financing channels remains one of the most significant barriers to their sustainable development. In the traditional banking system, strict collateral requirements, credit evaluation standards, and information asymmetry frequently hinder SMEs from securing formal credit, driving them to seek alternative and informal financing channels which may carry higher level of financial risks. Access to finance for SMEs is not uniform across regions. [4] demonstrate that during the financial crisis in the United Kingdom, SMEs in regions distant from bank headquarters experienced more significant credit constraints due to a "flight to headquarters" effect. Similarly, [5] identify a "low SME equity equilibrium" in Wales, characterized by both low demand and supply of equity, with a concentration of deals in urban centers like Cardiff. These studies underscore the spatial disparities in SME financing, justifying our geographical approach to examining shadow banking's role in SME survival.

In practice, many SMEs struggle to obtain loans from commercial banks, primarily due to inadequate collateral and limited financial transparency [3, 6]. To address short-term liquidity needs, these firms often rely on informal financing sources, including trade credit, private lending, and underground finance. Within this context, shadow banking has emerged as a crucial alternative financing mechanism [7]. Unlike traditional financial institutions, shadow banking entities operate with greater flexibility, more complex intermediation structures, and less stringent regulatory oversight, thus offering SMEs an important, though risky, channel for capital access.

The involvement of non-financial enterprises in shadow banking activities adds further complexity to the financing ecosystem. In particular, when non-financial listed companies engage in credit intermediation through entrusted loans, entrusted financing, or off-balance-sheet instruments, they contribute to the formation of non-traditional financial linkages within the corporate sector. These practices can partially alleviate the financing constraints faced by SMEs, especially as larger, state-affiliated firms typically have preferential access to formal banking credit [8]. As a result, non-financial enterprises may function as informal financial intermediaries, redistributing capital across enterprises in ways that are not captured by conventional financial intermediaries. While this type of credit intermediation could improve liquidity for constrained SMEs and enhance their survival prospects, it also introduces new systemic risks. The shadow banking system lacks formal risk assessment frameworks and prudential controls, and the credit it extended is often characterized by higher costs and shorter maturities. These factors can create adverse financial incentives and exacerbate the debt burden of SMEs, potentially increasing the likelihood of firm exit. Furthermore, due to the inter-enterprise nature of shadow credit flows, financial stress in one segment of the production system may be transmitted to other firms through capital linkages, thereby generating contagion effects within the regional economy.

This study addresses two key research questions: (1) How does shadow lending by non-financial enterprises influence the survival rates of SMEs within a region? (2) Are there spatial spillover effects through which shadow lending activities affect SME survival in neighboring regions? Following this, we examine the implications of shadow lending activities by non-financial enterprises on the survival dynamics of SMEs at the regional level. Specifically, we examine how informal credit flows originating from non-financial listed companies influence the mortality rate of SMEs within the same region. In addition, the study explores the presence of spatial spillover effects, assessing whether credit diffusion through shadow banking channels affects SME survival in neighboring regions via interregional financial linkages. To capture these dynamics, the study employs spatial econometric techniques, which are particularly well-suited to identify geographically distributed effects and interdependencies. By analyzing spatial correlation patterns in SME mortality, this paper uncovers how the informal financial system contributes not only to localized economic outcomes but also to broader systemic phenomena such as regional risk transmission, financial contagion, and structural vulnerability. The empirical results suggest that shadow banking provides vital liquidity support to financially constrained SMEs, particularly in regions underserved by formal financial institutions. While this financing channel appears to alleviate immediate funding pressures and enhance firm survival, the potential long-term implications remain ambiguous. Given the higher cost and lower regulatory oversight associated with shadow credit, concerns about financial fragility persist. Furthermore, the spatial spillover effects identified in the analysis imply that localized financial dynamics may propagate through interregional linkages, warranting further investigation into their systemic implications.

By integrating insights from the literature on financial intermediation, financial networks, and spatial spillover, this study contributes to a deeper understanding of the complex interdependencies between informal finance and enterprise dynamics. It emphasizes that shadow banking should be viewed not merely as a supplementary financing channel, but as a structurally embedded component of the production system with both stabilizing and destabilizing potential. Accordingly, the findings carry important policy implications for improving the governance of informal financial activity, managing systemic risk, and fostering the long-term resilience of SMEs in an increasingly networked financial environment.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature and develops the corresponding research hypotheses. Section 3 describes the dataset, which includes annual data from 2012 to 2022 on shadow lending activities by non-financial enterprises and the SME mortality across 243 prefecture-level cities in China. This section also outlines the data processing procedures, definitions of key variables, and the specification and diagnostic testing of the spatial econometric models employed. Section 4 presents the empirical analyses of the impact of non-financial enterprise shadow banking on regional SME mortality rates, employing the Spatial Durbin Model (SDM) as the primary estimation approach. Section 5 conducts robustness checks to verify the consistency of the findings. Section 6 provides additional analyses, including heterogeneity tests and an exploration of the spatial effects of policy shocks. Section 7 concludes the paper with a summary of key findings and policy implications.

2 Literature and hypotheses

From the perspective of financial intermediation theory, shadow banking functions as a non-traditional financial intermediary and plays a critical role in broadening credit supply and diversifying sources of financing [9]. In contrast to the traditional banking sector, which operates under stringent regulatory oversight and risk management protocols, shadow banking is characterized by greater operational flexibility. This flexibility enables it to provide supplementary liquidity during periods of credit contraction. Consequently, shadow banking has become particularly instrumental in supporting firms that are unable to secure financing through formal banking institutions, especially micro and small enterprises. These enterprises frequently face significant barriers to accessing affordable credit, and shadow banking serves as an alternative financing channel that effectively alleviates their funding constraints [7]. Recent evidence based on the TENET framework suggests that financial risk may spread more rapidly within shadow banking networks during tail events [10].

A substantial gap exists between the financing needs of SMEs and the supply capacity of the traditional financial system. Due to a lack of sufficient collateral and pronounced information asymmetries with banks, SMEs are often excluded from formal credit markets [2, 11]. As a result, many SMEs are compelled to rely on non-traditional financing channels, such as informal lending and trade credit, to sustain their daily operations. Shadow banking, functioning as a financing mechanism similar to trade credit, provides short-term liquidity to these firms, thereby helping them navigate temporary cash flow shortages and reducing the risk of business failure arising from disruptions in the capital chain. As a supplementary source of financing, shadow banking plays an especially important role when traditional financial institutions are unable to meet the funding needs of SMEs. It thus emerges as a critical alternative channel through which financially constrained firms can access much-needed capital [12, 13]. Network-based models show that incorporating inter-firm relationships can improve credit risk prediction for SMEs with limited data [14].

The diversification of financing channels plays a significant and positive role in the survival of SMEs. [15] emphasize that SMEs require different types of financing at various stages of their development, and that access to a diversified set of financing options is essential for their sustained growth. [3] further demonstrate that SMEs are subject to more severe financial and legal constraints compared to larger firms, and that the diversification of financing channels can help alleviate these constraints and facilitate enterprise growth. The effect of financing diversification on systemic risk depends on the structure of financial networks [16].

Although part of the credit demand of SMEs has been met, a substantial portion of latent demand remains unmet and has not been translated into actual financing. This further underscores the role of shadow banking in filling the gaps left by the traditional financial system by providing flexible financing mechanisms. Shadow banking can offer an alternative source of capital for SMEs, helping them overcome liquidity constraints and reducing the risk of bankruptcy caused by financing difficulties. Shadow banking activities exhibit significant regional disparities. [17] find that in China, the impact of collateral monetary policy on shadow banking by non-financial firms varies across regions, influenced by local financial development levels. [18] further reveal that regional advancements in financial technology are positively associated with increased shadow banking activities among non-financial firms. These findings highlight the necessity of considering regional contexts when analyzing shadow banking phenomena. In addition to regional factors, macro-level economic policies and regulations have significant effects on the growth of shadow banking. [19] argue that fintech innovations and regulatory arbitrage contribute to the rise of shadow banks, highlighting how less regulated financial activities can thrive in the gaps left by formal banking systems. Furthermore, [20] demonstrate that collateral monetary policy and its impact on shadow banking vary significantly across firms and regions, emphasizing the importance of considering broader financial stability measures when examining shadow banking dynamics.

Accordingly, existing literature suggests that when non-financial enterprises provide financing to SMEs through shadow banking channels, it may alleviate their funding constraints, enhance their capacity to survive, and ultimately reduce firm exit rates. Based on this reasoning, the following hypothesis is proposed.

H1a: Non-financial enterprises' shadow lending activities reduce local SMEs' mortality rate.

However, although shadow banking may to some extent alleviate short-term liquidity constraints and expand financing channels for SMEs, its high interest rates and elevated financing costs may also exert adverse effects on the long-term financial health of these firms. [21] points out that while informal financing can facilitate access to capital, it does not exhibit a significant positive correlation with long-term firm growth or productivity.

Furthermore, although the financing terms offered by shadow banking institutions are generally more flexible than those of traditional banks, they are often associated with higher interest rates and less transparent contractual conditions. [22] argue that while borrowing from multiple lenders may increase short-term liquidity, such dispersed lending relationships often result in higher overall financing costs. For SMEs that rely on shadow banking, the

inherent risks and opaque lending terms further exacerbate financial pressures and may undermine their long-term viability.

Moreover, the study by [23] indicates that although increased competition in the credit market may lead to lower lending rates, such rate reductions do not significantly expand the credit available to SMEs, particularly in the presence of persistent financing constraints. This suggests that even when shadow banking provides an alternative source of financing, firms may still face considerable financial burdens arising from high costs and elevated risks, which could ultimately increase their probability of default. Based on the above theoretical and empirical insights, it can be inferred that while shadow banking offers short-term liquidity support, its high-risk and high-cost nature may intensify financial stress for firms. This effect is likely to be more pronounced in unstable financial environments, potentially exacerbating the failure rates of SMEs. Therefore, the following hypothesis is proposed.

H1b: Non-financial enterprises' shadow lending activities improve local SMEs' mortality rate.

When examining the mortality of SMEs, it is essential to consider both region-specific factors and spatial interdependence, which refers to the mutual influences across geographic areas. The mortality of SMEs in one region may be shaped by economic activity in neighboring regions. The interregional flow of financial resources, spillover effects along industrial and supply chains, and the spatial diffusion of economic activity can all influence the business environment and the survival prospects of firms in adjacent areas [24]. Studies in spatial econometrics have shown that regional economic activities exhibit significant spatial spillover effects [25]. The survival or failure of firms is often affected by surrounding economic conditions. For example, [26] find that the emergence of new firms within a region can indirectly affect regional employment and economic development by enhancing competition and improving supply conditions.

Second, the flow of funds within the shadow banking system may also generate spatial spillover effects. Financing activities conducted by non-financial enterprises through shadow banking channels may not only affect SMEs within the local region but may also influence the financing environment and survival conditions of SMEs in neighboring areas through interregional financial linkages. Existing studies have shown that entrusted loans, as a component of shadow banking activities, can facilitate the cross-regional allocation of capital and thereby influence the distribution of financial resources and financing conditions across regions [27]. This effect is particularly pronounced in regions with a high degree of economic integration, where the flow and allocation of shadow banking capital may exert either positive or negative spillover effects on nearby SMEs. [21] find that the extent of dependence on shadow banking varies across regions, with firms in coastal areas being more reliant on shadow banking compared to those in central regions. Such disparities in capital flows may result in uneven financing capacities among SMEs across different regions, thereby influencing their mortality.

These findings suggest that the flow of shadow banking capital may produce spatial spillover effects through interregional financial interactions, ultimately shaping the survival prospects of SMEs. Therefore, it is necessary to account for the spatial correlation of SME mortality and the spatial spillover effects arising from non-financial enterprises' shadow lending activities. Based on the above analysis, the following hypothesis is proposed.

H2: The mortality of SMEs exhibits spatial dependence, and non-financial enterprises' shadow lending activities generate spatial spillover effects on SME mortality across regions.

TABLE 1 Summary statistics.

Variables	N	Mean	SD	Min	P50	Max
MSEMR	2,673	11.779	9.249	0.869	9.252	46.458
SL	2,673	15.845	9.806	0	20.596	29.387
SL_sum	2,673	16.583	9.408	0	20.808	29.395
PopDen	2,673	5.835	0.994	0.964	6.005	8.212
PerGDP	2,673	10.88	0.553	9.084	10.869	13.056
Edu	2,673	22.539	0.788	19.054	22.531	25.486
FD	2,673	26.031	1.171	23.379	25.819	29.898
Gov	2,673	0.475	0.218	0.058	0.447	1.541
Service	2,673	44.012	10.055	11.4	43.84	83.9

This table reports the summary statistics of the variables included in the empirical analyses over the period of 2012–2022. Within the table, variables are listed with their corresponding number of observations (N), mean, standard deviation (SD), and the values at the minimum (Min), median (P50), and maximum (Max). Regression t-statistics are reported in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively. See Table 1 for detailed definitions of these variable.

3 Empirical design

3.1 Sample data

To examine the impact of shadow lending of non-financial enterprises on the regional survival environment of SMEs, our study focuses on a sample of 243 prefecture-level cities in China over the period from 2012 to 2022. The empirical analysis relies on three main types of data. First, data on non-financial enterprises' shadow lending activities are obtained from the CNRDS database. Second, data used to measure SME mortality are derived from the industrial and commercial enterprise registration records available through AiQicha. Following the classification criteria proposed in the *Opinions of the State Administration for Industry and Commerce on Further Improving the Development of the Micro and Small Enterprise Directory* (Gongshanggezi [2015] No. 172), enterprises with registered capital of 5 million RMB or less are classified as micro and small enterprises. Third, data used to construct the spatial weight matrices and regional-level control variables are drawn from the *China Statistical Yearbook* and the CNRDS database. To minimize the influence of outliers, all continuous variables are winsorized at the 1% level. After data cleaning and processing, the final balanced panel includes 2,673 city-year observations. Descriptive statistics of all variables are presented in Table 1.

TABLE 2 Lagrange Multiplier Tests.

Models	LM tests		Robust LM tests	
	Test Statistic	P-value	Test Statistic	P-value
Spatial Error Model	2818.977	0.000	2129.215	0.000
Spatial Lag Model	793.643	0.000	103.882	0.000

This table reports the results of Lagrange Multiplier (LM) tests and Robust Lagrange Multiplier tests for spatial econometric models.

3.2 Testing the applicability of the spatial econometric model

Before selecting an appropriate spatial econometric model, it is standard practice to conduct a series of specification tests to assess model suitability. Commonly used spatial econometric models include the Spatial Lag Model (SLM), the Spatial Error Model (SEM), and the Spatial Durbin Model (SDM). To determine the most appropriate model for this study, this section first applies the Lagrange Multiplier (LM) test and the Robust LM test to the data. If the results indicate the presence of only spatial error dependence, the SEM is considered more appropriate. If only spatial lag dependence is significant, the SLM is preferred. However, if both effects are found to be statistically significant, the SDM is deemed the most suitable specification [28, 29]. The SDM allows for more complex spatial interactions, making it more suitable for analyzing spatially dependent data in our study.

According to the LM and Robust LM test results presented in Table 2, both spatial error dependence and spatial lag dependence are statistically significant ($P < 0.01$), indicating that the SDM is more appropriate for the analysis in this study. Subsequently, a Hausman test is conducted to determine whether to adopt a random effect or fixed effect specification. The test yields a statistic of 29.1858 with a p-value of 0.0001, leading to the rejection of the null hypothesis. Therefore, the fixed effects SDM is selected. Finally, post-estimation diagnostic tests including the Wald test and the Likelihood Ratio (LR) test are employed to examine whether the SDM can be simplified to either the SLM or SEM specification.

As shown in Table 3, the results of the LR tests reject the null hypothesis, indicating that the two-way fixed effects model is more appropriate. The Wald test further confirms the significance of spatial lag effects and suggests that there is no simple linear relationship between the spatial lag term and the spatial error term. Therefore, the SDM is better suited to capture spatial dependence and spatial spillover effects, making it a more appropriate choice for the empirical analysis in this study.

3.3 Specification of the spatial durbin model and variable definitions

Given the complexity of capital flows involved in the non-financial enterprises' shadow lending activities, and the difficulty in tracing the destination of such funds at the individual level within the sample of listed firms, it is plausible that these activities generate spatial spillover effects on SMEs in neighboring regions. To capture these cross-regional financial interactions and economic fluctuations, we constructs a spatial weight matrix to reflect the spatial interdependence across regions. Based on the applicability tests conducted in the previous section, the SDM is selected as the appropriate framework for analyzing both the direct impact of shadow banking by non-financial enterprises on regional SME mortality and its potential spillover effects on adjacent areas. The model is specified in Equation 1:

$$MSEMR_{it} = \alpha + \rho \sum_{j=1}^{n} W_{ij} MSEMR_{jt} + \beta_1 SL_{it} + \beta_2 Controls_{it} + \theta_1 \sum_{j=1}^{n} W_{ij} SL_{jt} + \theta_2 \sum_{j=1}^{n} W_{ij} Controls_{jt} + City_i + Year_t + \varepsilon_{it} \quad (1)$$

where i denotes the prefecture-level city, and t denotes the year. W_{ij} represents the element of the spatial weight matrix, which captures the spatial relationship between city i and city j. We constructs three types of spatial weight matrices: a distance matrix ($W1$), an adjacency matrix ($W2$), and an economic matrix ($W3$). The dependent variable, $MSEMR_{it}$, refers to the mortality of micro and small enterprises in city i during year t. The key explanatory variable, SL_{it}, measures the amount of shadow lending by non-financial listed companies in city i at time t, following the calculation approach described in the previous sections. The spatially lagged dependent variable, $W_{ij}MSEMR_{jt}$, captures the impact of SME mortality in neighboring cities on that of city i. The term $W_{ij}SL_{jt}$ represents the spatial lag of shadow banking activities, indicating the influence of shadow banking intensity in adjacent regions on SME mortality in city i.The vector $Controls_{it}$ includes a set of control variables for city i at time t, while $W_{ij}Controls_{jt}$ denotes the spatially lagged controls. $City_i$ and $Year_t$ represent city and year fixed effects, respectively. α is the intercept term, and ε_{it} is the error term.

To mitigate the influence of spurious correlations on the empirical results, we select a set of regional-level control variables. Specifically, the following variables that may affect the volume of credit disputes are included: population density (*PopDen*), economic development level (*PerGDP*), education level (*Edu*), financial development (*FD*), government support (*Gov*), and industrial structure (*Service*). The definitions and calculation methods of these variables are provided in Table 4. All variables are measured at year t.

In conventional non-spatial regression models, the magnitude of the regression coefficients can be directly interpreted as the marginal effects of explanatory variables on the dependent variable. However, as discussed earlier, due to the presence of spatial dependence in spatial econometric models, explanatory variables in a given region may not only influence the local dependent variable but also exert indirect effects on neighboring regions through spatial interaction mechanisms. Therefore, once spatial lag terms are included in the model, the estimated coefficients cannot be directly interpreted as the total effects of the explanatory variables on the dependent variable [28, 29]. To address this issue, spatial effects are decomposed following the methodology proposed by [28]. A significantly positive direct effect indicates that non-financial enterprises' shadow lending activities increase the mortality of SMEs within a region, thereby supporting Hypothesis H1b. Conversely, a significantly negative direct effect implies that such activities reduce SME mortality, supporting Hypothesis H1a.

TABLE 3 Likelihood Ratio and Wald Tests.

Tests	Null hypothesis *H*0	Test statistic	P-value
LR Tests	Spatial fixed effects model is nested within the two-way fixed effects model	121.32	0.000
	Time fixed effects model is nested within the two-way fixed effects model	1900.64	0.000
Wald Tests	All coefficients of spatially lagged explanatory variables are equal to zero (no spatial lag effects)	138.28	0.000
	Existing a specific linear relationship between the spatial lag and spatial error terms	159.56	0.000

This table reports the results of Likelihood Ratio (LR) and Wald Tests for the Spatial Durbin Model (SDM).

TABLE 4 Definitions of Variables.

Variables	Definitions	Calculations
MSEMR	Micro and Small Enterprise Mortality Rate	The number of exited firms divided by the sum of firms existing at the end of the previous year and newly established firms in the current year, multiplied by 100%
SL	Shadow Lending	Natural logarithm of the sum of entrusted loans, entrusted wealth management products, and private loans (i.e., other receivables) financed by the listed companies registered within the city
SL_sum	Sum of *SL* and chain-based shadow lending activities	The natural logarithm of the scale of non-financial listed firms involves in intermediary (i.e., *SL*) and chain-based shadow banking activities as lenders, where chain-based shadow banking activities refers to the amount of listed companies indirectly involving in shadow banking activities through purchasing trust products, bank wealth management products, and internet financial products
PopDen	Population Density	the natural logarithm of the number of persons per square kilometer in each region
PerGDP	Per capita GDP	The natural logarithm of the GDP *per capita*
Edu	Education	The natural logarithm of education expenditure in each region (in yuan)
FD	Financial Development	Degree of financial development, measured as the ratio of bank loan balances to local GDP.
Gov	Government Support	The ratio of public fiscal revenue (in yuan) to public fiscal expenditure (in yuan)
Service	Proportion of Tertiary Industry	Share of tertiary industry in GDP (%)

This table reports the definition, calculation (if applicable) of main variables used in our empirical analyses.

If the spatial autoregressive coefficient ρ is significantly different from zero, it indicates the presence of spatial dependence in SME mortality, meaning that the survival environment of SMEs in one region is significantly influenced by that of neighboring regions. Furthermore, if the spatially lagged explanatory variable θ is significantly different from zero, it suggests that shadow banking activities have cross-regional spillover effects. In other words, SL_{it} affects not only the SME mortality in its own region but also that of surrounding or economically linked regions through spatial interaction, thereby validating Hypothesis H2.

4 Empirical results

4.1 Test for spatial correlation

Before implementing Spatial Durbin Model (SDM), we conduct spatial correlation tests using the Global Moran's I Index [30] and the Local Moran's I Index [31] as part of a spatial exploratory analysis. These tests assess whether regional credit risks exhibit spatial dependence, justifying the use of spatial econometric models in subsequent analyses. The Global Moran's I Index is used to test

global spatial correlation, measuring the overall spatial distribution of regional credit risks. The formula for the Global Moran's I Index is as follows:

$$I = \frac{N\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{W\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2}$$

where in Equation 2, $N = 243$, representing the number of prefecture-level cities in the sample. x_i and x_j denote credit risk values for cities i and j respectively. $\bar{x}$ is the mean valure of x_i and x_j, and w_{ij} is the spatial weight matrix defining the relationship between city i and j. W is the sum of w_{ij}. The Global Moran's I Index typically ranges between −1 and 1 ($I \in [-1, 1]$). A positive value of I indicates positive spatial correlation, meaning that neighboring cities exhibit similar credit risk characteristics. The larger the absolute value of I, the stronger the spatial correlation. Conversely, a negative value of I indicates negative spatial correlation, meaning that neighboring cities exhibit dissimilar credit risk characteristics. If I equals 0, it suggests that the credit risks of cities are independently distributed in space, with no spatial correlation.

The significance of the spatial correlation is tested using the standardized Z-statistic, calculated as:

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} \tag{3}$$

where in Equation 3, I is the Global Moran's I value calculated from the formula above. $E(I)$ is the expected value of I, typically equal to $-1/(N-1)$, $Var(I)$ is the variance of the Global Moran's I. Table 5 presents the Global Moran's I indices for SMEs mortality under three different spatial weight matrix specifications. The results indicate that the Global Moran's I values are positive across all three spatial weight matrices and that the null hypothesis of no spatial correlation is rejected at the 1% significance level. This suggests that SME mortality rate exhibits significant spatial autocorrelation and spatial dependence among prefecture-level cities in China during the sample period.

In other words, if the business environment for SMEs is favorable in a given city, neighboring regions also tend to exhibit similarly favorable conditions for SME survival, and *vice versa*. This spatial clustering effect implies that the analysis of SME mortality should not be confined to individual cities but should instead account for cross-regional transmission mechanisms and interdependencies. Therefore, based on the results of the spatial correlation tests, it is both reasonable and necessary to adopt a spatial econometric approach to examine the impact of shadow lending of non-financial enterprises on SME mortality.

To further examine the spatial spillover effects of regional credit risks, we compute the Local Moran's I Index for each city, assessing the statistical significance of spatial clustering patterns [31]. The Local Moran's I Index is given by (see Equation 4):

$$I_i = \frac{(x_i - \bar{x})\sum_{j=1}^{N} w_{ij}(x_j - \bar{x})}{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}} \tag{4}$$

Using a distance-based spatial weight matrix (w_1), we compute the Local Moran's I index for all 243 prefecture-level cities over time. Figure 1 presents the Moran's I scatter plots for selected years (2012, 2015, 2017, and 2022). The x-axis represents the standardized Z-statistic for credit risk in each city, while the y-axis represents the spatially lagged credit risk in neighboring cities (Wz). The scatter plots illustrate that most cities fall within the first and third quadrants, indicating a strong spatial clustering effect where high-credit-risk cities are located near other high-risk cities, and low-risk cities are clustered together. These findings confirm that regional credit disputes do not occur in isolation but propagate through spatial financial linkages, supporting the need for spatial econometric modelling in our study.

4.2 Analysis of total effects

Based on Equation 1, this section analyzes the total effects of shadow lending of non-financial enterprises on the survival of SMEs, covering both local and neighboring regions. The estimation results of the SDM under three spatial weight matrices are presented in Table 6. The distance matrix ($W1$) serves as the benchmark, while the adjacency matrix ($W2$) and the economic matrix ($W3$) are used for robustness checks. The table reports the estimation results for both random effects and fixed effects specifications of the SDM.

As shown in Table 6, the estimation results from the SDM under all three spatial weight matrices indicate that the null hypothesis is rejected at the 1% significance level in the Hausman test. This finding is consistent with the results reported in Section 3.1, suggesting that the fixed effects SDM provides a better fit for the data and is therefore adopted for interpretation. Furthermore, the coefficient of the spatially lagged dependent variable (ρ) is significantly positive at the 1% level across all specifications, indicating strong spatial dependence among the variables. This result is in line with the Global Moran's I test presented in Section 4.1, confirming that the survival environments of SMEs in prefecture-level cities are spatially interrelated. In other words, the business environment for SMEs in one city significantly influences that of neighboring or proximate cities under spatial weighting structures.

With respect to the core explanatory variable, the coefficient of shadow banking by non-financial enterprises (SL) is significantly negative at the 1% level across all six regressions using the three spatial weight matrices. Its spatially lagged counterpart (WSL) is also significantly negative in all specifications, with the exception of the fixed effects model under the adjacency matrix ($W2$), where it is significant at the 5% level ($t = -2.21$); in the remaining cases, the coefficient remains significantly negative at the 1% level. These results provide preliminary evidence that shadow lending of non-financial enterprises reduce SME mortality, thereby having a positive impact on SME survival, which is consistent with some prior research [12, 13]. Moreover, the negative and significant coefficient on the spatial lag of shadow banking suggests the presence of a spillover effect, whereby shadow banking activity in neighboring regions further contributes to a reduction in local SME mortality. These findings offer initial empirical support for Hypotheses H1a and H2, indicating that non-financial enterprise shadow banking mitigates regional SME mortality and exerts significant spatial spillover effects on the SME survival environment.

10.3389/fphy.2025.1612653

TABLE 5 Moran's I Index.

Year	Moran's I (W1)	Z-value	Moran's I (W2)	Z-value	Moran's I (W3)	Z-value
2012	0.067	10.112	0.304	6.962	0.269	8.665
2013	0.058	8.936	0.249	5.855	0.218	7.201
2014	0.038	7.126	0.160	4.476	0.140	5.497
2015	0.069	10.129	0.310	6.916	0.261	8.205
2016	0.073	10.252	0.365	7.817	0.300	9.067
2017	0.076	10.573	0.343	7.328	0.281	8.439
2018	0.064	9.013	0.335	7.174	0.255	7.720
2019	0.063	8.900	0.342	7.303	0.236	7.129
2020	0.050	7.174	0.262	5.634	0.159	4.838
2021	0.064	9.226	0.352	7.658	0.151	4.681
2022	0.063	9.205	0.277	6.171	0.187	5.889

This table reports the Global Moran's I Index for mortality micro and small enterprises mortality amounts across prefecture-level.

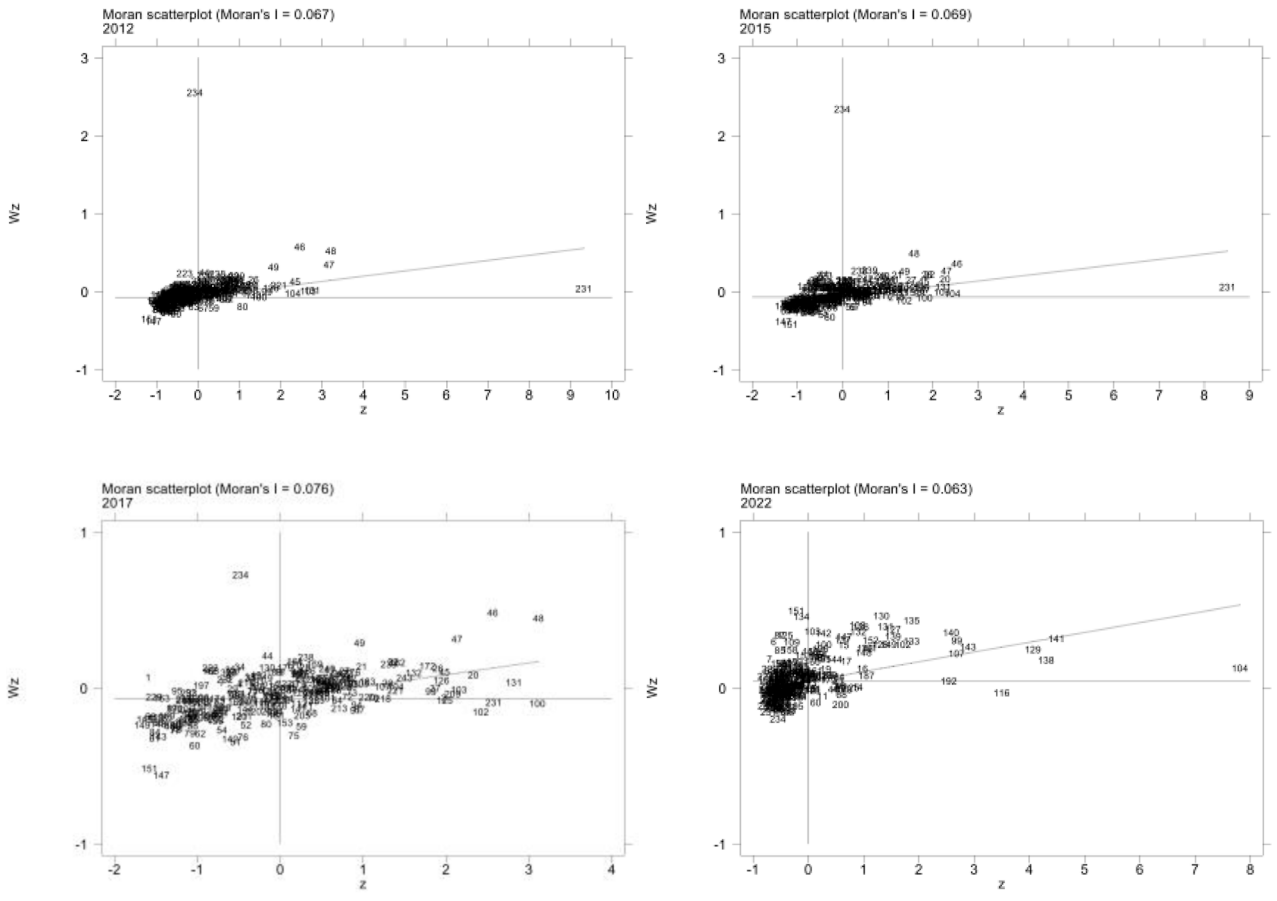


FIGURE 1
Moran's I scatter plots of regional credit risk.

TABLE 6 Spatial Durbin Model (SDM) Regression Results of shadow lending and SMEs mortality rate.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Distance matrix (*W1*)		Adjacency matrix (*W2*)		Economic matrix (*W3*)	
	Random effect	Fixed effect	Random effect	Fixed effect	Random effect	Fixed effect
SL	−0.033***	−0.040***	−0.041***	−0.029**	−0.053***	−0.047**
	(-2.58)	(-3.22)	(-3.31)	(-2.53)	(-2.70)	(-2.47)
WSL	−0.201***	−0.959***	−0.076***	−0.048**	−0.420***	−0.399***
	(-3.41)	(-6.54)	(-3.56)	(-2.21)	(-8.96)	(-7.82)
PopDen	0.120	1.003**	1.333***	0.962**	1.773***	3.231***
	(0.39)	(2.33)	(4.05)	(2.35)	(3.76)	(4.94)
PerGDP	2.697***	2.897***	−1.272**	2.373***	1.776**	4.925***
	(4.67)	(4.10)	(-2.40)	(3.61)	(2.17)	(5.02)
Edu	0.333	0.178	0.065	−0.515	−1.141	−2.045**
	(0.66)	(0.26)	(0.14)	(-0.83)	(-1.59)	(-2.01)
FD	−0.688*	−0.563	−1.787***	−1.064**	−2.085***	−0.562
	(-1.69)	(-1.00)	(-4.58)	(-2.05)	(-3.58)	(-0.66)
Gov	−1.364	1.203	9.426***	0.677	4.442**	−1.593
	(-0.98)	(0.78)	(7.22)	(0.48)	(2.20)	(-0.72)
Service	0.020	−0.012	−0.072***	−0.022	−0.048*	−0.009
	(1.02)	(-0.57)	(-3.75)	(-1.13)	(-1.65)	(-0.29)
ρ	0.945***	0.894***	0.626***	0.492***	0.367***	0.171***
	(57.85)	(29.14)	(42.31)	(26.13)	(18.68)	(7.89)
σ^2	17.360***	14.951***	16.721***	12.853***	42.718***	34.686***
	(34.65)	(36.57)	(33.67)	(35.70)	(34.56)	(36.54)
Constant	−55.410		65.410***		109.220***	
	(-1.56)		(9.02)		(7.20)	
WControls	Yes	Yes	Yes	Yes	Yes	Yes
City FE	No	Yes	No	Yes	No	Yes
Year FE	No	Yes	No	Yes	No	Yes
Obs	2,673	2,673	2,673	2,673	2,673	2,673
R^2	0.486	0.502	0.457	0.012	0.380	0.131
Hausman Test	29.1858		1149.028		614.4279	
	P-value = 0.0001		P-value = 0		P-value = 0	

This table reports the results of our main regression model of Equation 1. See Table 3 for detailed variable definitions. The dependent variable is the micro and small enterprises mortality rate (*MSEMR*), while the independent variable is shadow lending by local non-financial listed firms (*SL*).The results are reported under different spatial weighting matrices: distance matrix [Column (1)–(2)], adjacency matrix [Column (3)–(4)], and economic matrix [Column (5)–(6)], each under both random and fixed effects models. Two-way fixed effects are included where applicable. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

4.3 Analysis of decomposed effects

As discussed in Section 3.2, regional credit risk exhibits spatial contagion. When shadow lending by non-financial enterprises in one city affects the credit risk in other cities, the resulting spillover effects may in turn influence the originating city. Therefore, although the regression coefficients reported in Table 6 provide useful insights, they cannot be directly interpreted as the total or spillover effects of the explanatory variables on the dependent variable. To more accurately assess the spatial spillover effects of shadow lending of non-financial enterprises on both local and neighboring SME mortality, this section adopts the methodological framework proposed by [28]. Specifically, the SDM estimation results under the two-way fixed effects specification reported in Table 6, including Models (2), (4), and (6), are decomposed into direct effects, indirect effects, and total effects. The decomposition results for each variable are presented in Table 7.

Regarding the direct effects, the non-financial enterprises' shadow lending activities exhibit significantly negative direct effects under all three spatial weight matrices ($W1$, $W2$, and $W3$), with coefficients of −0.088, −0.039, and −0.054 respectively, all statistically significant at the 1 percent level. This indicates that shadow lending of non-financial enterprises have a favorable direct impact on the local survival environment of SMEs, providing further support for Hypothesis H1a, which posits that such activities reduce local SME mortality. Among the control variables, population density (*PopDen*) and economic development (*PerGDP*) show predominantly positive direct effects, suggesting that higher levels of economic activity and population density may intensify market competition, thereby increasing pressure on SME survival. Turning to the indirect effects, the spatial spillover effects of shadow lending (*SL*) are also significantly negative across all three spatial weight matrices. Notably, under the distance matrix ($W1$), the coefficient is −10.911 with a t-value of −2.04, indicating that shadow banking activities in neighboring regions exert a statistically significant spillover effect by reducing local SME mortality. This finding supports Hypothesis H2, which states that shadow banking generates significant spatial spillover effects.

Interestingly, among the control variables, educational expenditure (*Edu*) and financial development (*FD*) exhibit significantly positive indirect effects. This suggests that, through mechanisms such as intensified interregional competition, resource reallocation, and uneven financial access, these factors may inadvertently increase SME mortality in neighboring regions. While typically associated with economic growth, in this context, such variables may raise survival pressure on SMEs. For instance, a financial system biased toward large enterprises may increase financing costs for smaller firms, or a concentration of public resources in more developed areas may weaken the competitiveness of less developed regions, ultimately affecting the survival of local SMEs.

Regarding the total effects, shadow banking by non-financial enterprises exhibits significantly negative total effects under all three spatial weight matrix specifications. This finding further confirms that shadow banking not only directly improves the local survival environment of SMEs but also benefits SMEs in neighboring regions through spatial spillover effects. Among the control variables, the significantly positive total effects of educational expenditure and financial development reinforce their role in increasing the survival pressure faced by SMEs.

In summary, both the direct and indirect effects of shadow banking by non-financial enterprises are significantly negative. This indicates that such activities not only enhance the survival environment for local SMEs but also improve the survival conditions of SMEs in neighboring regions through spatial spillovers. These results suggest that shadow lending helps alleviate financing constraints faced by SMEs, strengthens their resilience against risk, and enhances their competitiveness in the market. The shadow banking system provides a more flexible source of financing for these firms, especially in contexts where formal financial institutions are unable to fully meet their funding needs. As a result, shadow banking contributes to extending the life cycle of SMEs and reducing their mortality. Our findings align with prior research indicating that shadow banking activities can alleviate financing constraints for SMEs, thereby reducing their mortality rates [7, 12]. Additionally, the observed spatial spillover effects are consistent with studies highlighting the role of informal financial channels in facilitating inter-regional capital flows and influencing neighboring regions' economic dynamics [26, 27].

From a spatial perspective, the flow of capital and the allocation of financial resources through shadow banking are not confined to individual regions but can cross administrative boundaries, producing significant spatial spillover effects. This reinforces the role of shadow banking in supporting the continuity and development of SMEs. The empirical findings validate Hypotheses H1a and H2, indicating that shadow banking by non-financial enterprises significantly reduces SME mortality within a given region and simultaneously enhances the survival environment of SMEs in adjacent areas through spillover mechanisms. Furthermore, the statistical significance of the distance, adjacency, and economic matrices suggests that the influence of shadow banking activities on regional economies is both extensive and complex. Shadow banking not only plays a positive role in improving the survival prospects of SMEs in the originating region but also generates meaningful indirect effects on neighboring regions through spatial network interactions. Overall, shadow banking has significant direct and indirect effects on SME survival, and these effects extend beyond geographic boundaries.

5 Robustness tests

To ensure the validity of the empirical findings, several robustness checks are conducted in this study. First, by using the distance matrix ($W1$) as the baseline spatial weight matrix and incorporating the adjacency matrix ($W2$) and the economic matrix ($W3$) as alternative specifications, the sensitivity of the results to different spatial structures can be evaluated. The corresponding estimation results have been presented and analyzed in Section 4.

Next, to further test the robustness of the results, the explanatory variable is replaced with the natural logarithm of the sum of credit intermediation-type and credit chain-type shadow banking activities (*SL_sum*) to capture the scale of shadow banking by non-financial enterprises more comprehensively. The SDM with two-way fixed effects is estimated, and the results

TABLE 7 Spatial Spillover Effects—Decomposition of Effects.

Variables		(1)	(2)	(3)
		Distance matrix (*W1*)	Adjacency matrix (*W2*)	Economic matrix (*W3*)
Direct effect	*SL*	−0.088***	−0.039***	−0.054***
		(-3.34)	(-3.02)	(-2.85)
	PopDen	0.672	0.691*	3.199***
		(1.64)	(1.94)	(5.80)
	PerGDP	2.765***	2.463***	5.038***
		(3.88)	(3.61)	(4.86)
	Edu	0.971	0.149	−1.779
		(1.17)	(0.21)	(-1.61)
	FD	0.462	−0.568	−0.661
		(0.52)	(-0.89)	(-0.68)
	Gov	1.568	0.929	−1.573
		(1.11)	(0.68)	(-0.74)
	Service	0.018	−0.009	−0.005
		(0.79)	(-0.49)	(-0.17)
Indirect effect	*SL*	−10.911**	−0.115***	−0.494***
		(-2.04)	(-3.12)	(-8.61)
	PopDen	−64.523	−2.518**	2.320*
		(-1.20)	(-2.54)	(1.85)
	PerGDP	−41.614	0.466	2.647
		(-0.74)	(0.32)	(1.60)
	Edu	165.659*	6.794***	12.311***
		(1.69)	(4.01)	(5.44)
	FD	257.051*	6.727***	5.608**
		(1.88)	(4.25)	(2.53)
	Gov	69.135	2.263	−3.315
		(0.65)	(0.71)	(-0.83)
	Service	6.468**	0.133***	0.180***
		(2.13)	(3.47)	(3.43)
Total effect	*SL*	−10.999**	−0.154***	−0.548***
		(-2.04)	(-3.40)	(-8.62)
	PopDen	−63.852	−1.827	5.519***

(Continued on the following page)

TABLE 7 (*Continued*) Spatial Spillover Effects—Decomposition of Effects.

Variables		(1)	(2)	(3)
		Distance matrix (*W*1)	Adjacency matrix (*W*2)	Economic matrix (*W*3)
		(-1.18)	(-1.62)	(4.21)
	PerGDP	−38.849	2.930*	7.685***
		(-0.69)	(1.86)	(4.56)
	Edu	166.630*	6.942***	10.533***
		(1.70)	(3.52)	(4.55)
	FD	257.513*	6.159***	4.947**
		(1.88)	(3.19)	(2.03)
	Gov	70.703	3.192	−4.888
		(0.66)	(0.87)	(-1.14)
	Service	6.485**	0.124***	0.175***
		(2.13)	(2.82)	(3.40)

This table decomposes the spatial spillover effects into direct, indirect, and total effects under different spatial weighting matrices: distance matrix, adjacency matrix, and economic matrix. Two-way fixed effects are included. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

are reported in Table 8. The estimated coefficients of the key variable remain consistent in sign with those from the baseline regressions in Tables 6, 7. Across all specifications under the three spatial weight matrices, the direct effects of shadow banking by non-financial enterprises are significantly negative at the 1 percent level. This confirms that, regardless of the spatial weight structure employed, shadow banking consistently improves local credit conditions and reduces SME mortality, thereby supporting Hypothesis H1a.

The results for the indirect and total effects further demonstrate that the expanded measure of shadow banking activity (*SL_sum*), which encompasses a broader scope of financial operations, has a significantly negative effect at the 1 percent level across all spatial weight matrices. This indicates that shadow banking by non-financial enterprises affects not only local SME survival but also exerts spatial spillover effects on neighboring regions. These findings provide additional empirical support for Hypothesis H2.

In addition, to further examine the long-term impact of shadow lending of non-financial enterprises on the survival of SMEs, a two-way fixed effects panel regression is employed to assess the effect of shadow banking on local SME mortality in periods $t+1$ and $t+2$. The results of the robustness check for lagged effects are presented in Table 9.

The regression coefficients of shadow lending are consistently and significantly negative across the models estimating its impact on SME mortality in periods t, $t+1$, and $t+2$, with significance levels at the 1 percent or 5 percent thresholds. These results indicate that the non-financial enterprises' shadow lending activities have a significant and persistent effect in reducing SME mortality over time. This finding provides empirical support for the view that shadow banking alleviates financing constraints and enhances the long-term survival capacity of local SMEs. Overall, the results from Table 9 suggest that the mitigating effect of shadow banking on local SME mortality is not only significant in the current period but also persists into the subsequent 2 years. However, the effect becomes statistically insignificant beyond period $t+2$. These findings further confirm the robustness of the long-term positive impact of shadow banking on SME survival.

6 Further analysis

6.1 Regional heterogeneity analyses

To examine whether the impact of shadow banking by non-financial enterprises on regional SME mortality varies across different regions, we further divide the sample of 243 cities into two groups based on the "Hu Huanyong Line"[2] ("Hu line"), which geographically separates China into the southeastern and northwestern regions. The southeastern region is characterized by higher population density, greater urbanization, and more dynamic financial markets and business activities. In contrast, the northwestern region features lower population

2 The Hu Huanyong Line, first proposed by Chinese demographer Hu Huanyong in 1935, is a demographic and economic dividing line that runs diagonally from Heihe in Heilongjiang province to Tengchong in Yunnan province. The southeastern side of the line contains approximately 94% of China's population and economic activity, while the northwestern side is sparsely populated and less economically developed [34].

10.3389/fphy.2025.1612653

TABLE 8 Robustness tests: alternative measurements of independent variables.

Variables	(1)	(2)	(3)
	Distance matrix (*W1*)	Adjacency matrix (*W2*)	Economic matrix (*W3*)
SL_sum	−0.064***	−0.054***	−0.053***
	(-3.23)	(-2.74)	(-2.70)
WSL_sum	−1.775***	−0.176***	−0.371***
	(-7.76)	(-4.73)	(-6.63)
PopDen	3.450***	3.186***	3.193***
	(5.20)	(4.70)	(4.87)
PerGDP	5.707***	6.016***	4.854***
	(5.24)	(5.54)	(4.94)
Edu	−1.520	−1.594	−2.109**
	(-1.46)	(-1.55)	(-2.07)
FD	−0.265	−0.112	−0.595
	(-0.31)	(-0.13)	(-0.70)
Gov	−0.359	−1.891	−1.750
	(-0.15)	(-0.81)	(-0.78)
Service	−0.008	−0.003	−0.009
	(-0.25)	(-0.09)	(-0.30)
ρ	0.461***	0.191***	0.166***
	(6.03)	(9.90)	(7.54)
σ^2	35.593***	35.256***	34.896***
	(36.34)	(36.50)	(36.54)
WControls	Yes	Yes	Yes
Direct effect	−0.077***	−0.064***	−0.060***
	(-3.81)	(-3.16)	(-3.01)
Indirect effect	−3.459***	−0.221***	−0.459***
	(-5.01)	(-5.58)	(-7.49)
Total effect	−3.536***	−0.285***	−0.519***
	(-5.09)	(-5.82)	(-7.66)
City FE	Yes	Yes	Yes

(Continued on the following page)

TABLE 8 (*Continued*) Robustness tests: alternative measurements of independent variables.

	(1)	(2)	(3)
	Distance matrix (*W*1)	Adjacency matrix (*W*2)	Economic matrix (*W*3)
Year FE	Yes	Yes	Yes
Obs	2,673	2,673	2,673
R^2	0.358	0.011	0.135

This table reports the results of regressions including alternative proxies for the independent variables under different spatial weighting matrices: distance matrix, adjacency matrix, and economic matrix. The independent variable is replaced by *SL_sum*, refers to the sum of *SL*, and the amount that listed companies indirectly involving in shadow banking activities through purchasing trust products, bank WMPs, and internet financial products. Two-way fixed effects are included. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

TABLE 9 Robustness tests: the time-lagged effects.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	$MSEMR_t$		$MSEMR_{t+1}$		$MSEMR_{t+2}$	
SL	−0.035**	−0.041***	−0.042***	−0.049***	−0.027**	−0.031**
	(-2.44)	(-2.92)	(-3.01)	(-3.66)	(-2.11)	(-2.54)
PopDen		0.552		−0.012		0.254
		(1.16)		(-0.03)		(0.68)
PerGDP		4.721***		6.167***		5.610***
		(7.26)		(10.08)		(9.77)
Edu		2.623***		1.990***		0.623
		(3.71)		(3.02)		(1.01)
FD		1.222**		0.509		0.191
		(1.99)		(0.87)		(0.35)
Gov		−0.956		−1.922		−1.366
		(-0.60)		(-1.27)		(-0.97)
Service		0.059***		0.035*		−0.015
		(2.89)		(1.85)		(-0.85)
Constant	22.877***	−121.424***	20.782***	−101.835***	17.685***	−60.386***
	(72.66)	(-6.57)	(73.00)	(-5.78)	(71.40)	(-3.61)
City FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Obs	2,673	2,673	2,430	2,430	2,187	2,187
City	243	243	243	243	243	243
R^2	0.703	0.717	0.707	0.726	0.738	0.738

This table presents the time-lagged effects of shadow lending on the micro and small enterprises mortality rate across contemporaneous and lagged periods (t, t+1, and t+2). The panel regression models include two-way fixed effects. Standard errors clustered at the firm level. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

density, slower economic development, and relatively unbalanced financial growth.

As shown in the subgroup regression results in Table 10, the impact of shadow banking by non-financial enterprises on SME mortality is more pronounced in southeastern regions, while it is relatively insignificant in northwestern regions. In the southeast, the coefficient of shadow lending (*SL_sum*) is significantly negative under both the distance matrix (*W*1) and the adjacency matrix (*W*2), suggesting that shadow banking activities contribute to reducing SME mortality and generate positive spillover effects across neighboring areas. In contrast, although the coefficients of shadow lending in northwestern cities are also negative, most regressions fail to reach statistical significance. This reflects the weaker effectiveness of shadow banking in improving the survival environment of SMEs in less developed financial markets with lower capital mobility, such as those in the northwest.

Moreover, the spatial lag term of shadow lending (*WSL_sum*) is significantly negative in both the southeastern and northwestern regions, indicating that shadow banking activities produce positive spillover effects across regions. In the northwest, these spatial lag effects are significant under the geographic matrices (*W*1 and *W*2), suggesting that when local capital is limited, shadow banking can influence SME survival through spillovers from adjacent areas. However, under the economic matrix (*W*3), the effect is not significant, indicating that economically similar cities do not exhibit strong financial transmission effects in the northwest. By contrast, the southeastern region shows significant negative effects in both the direct and spatial lag terms, highlighting the important role of shadow banking in supporting SME survival in more developed economic environments. Using the Hu Line as a basis for regional division, this heterogeneity analysis further validates that the effects of shadow banking on SME mortality and its spatial spillovers vary significantly across regions.

6.2 The impact of shadow banking by non-financial enterprises on regional firm survival

To further investigate the impact of shadow lending of non-financial enterprises on the regional business environment and its spatial spillover effects, this section expands the sample to include all enterprises. The estimation results of the SDM with two-way fixed effects under the three spatial weight matrix specifications are reported in Table 11.

The regression results indicate that shadow lending of non-financial enterprises continue to have a statistically significant effect on the mortality of all firms. The coefficient for shadow lending is negative and significant at the 5 percent level under all three spatial weight matrix specifications, suggesting that shadow banking significantly reduces overall firm mortality. This finding is consistent with the earlier results based on the SMEs subsample and further supports the positive role of shadow banking in enhancing firm survival and reducing firm exits. However, the decomposition of spatial effects reveals that the impact of shadow banking on overall firm mortality resembles that of SMEs only under the distance matrix (*W*1). Under the alternative spatial weight specifications, the influence of shadow lending on overall firm mortality is comparatively weaker.

Overall, the effect of shadow banking by non-financial enterprises on the survival of all firms appears to be less pronounced than its effect on SMEs. Both the direct effects on local firms and the indirect effects on firms in neighboring regions are relatively limited in magnitude.

6.3 Mechanism analysis: the impact of shadow banking by non-financial enterprises on the survival of micro and small enterprises

This section employs a difference-in-differences (DID) approach to evaluate the impact of the 2018 *New Asset Management Regulations*[3], which sought to curb systemic financial risks by regulating shadow banking activities. The policy serves as an exogenous shock, allowing us to causally evaluate its effectiveness in reducing regional credit risks associated with shadow lending by non-financial enterprises.

As an informal financing channel, shadow banking serves as an important source of capital for SMEs that face difficulties in accessing traditional credit. However, the tightening of financial regulation directly affects the scale and liquidity of shadow banking. In particular, the implementation of the *New Asset Management Rules* had a direct impact on both the volume and flow of funds within the shadow banking system. While the primary objective of the policy tightening was to curb high-risk activities in shadow banking, which may have constrained credit supply in the short term, it also aimed to enhance long-term financial stability by reducing systemic risk through stricter oversight of risk-prone financial operations. To explore the underlying mechanism in greater depth, this section adopts a difference-in-differences (DID) approach combined with the Spatial Durbin Model (SDM) to examine the effects of the New AM Rules on shadow banking activities and their subsequent influence on SME survival. The SDM-DID model is specified as shown in Equation 5:

$$\begin{aligned} SMEMR_{it} = {} & \alpha + \rho \sum_{j=1}^{n} W_{ij} SMEMR_{jt} + \beta_1 \left(Treat_i \times Post_t \right) \\ & + \beta_2 Controls_{it} + \theta_1 \sum_{j=1}^{n} W_{ij} \left(Treat_i \times Post_t \right) \\ & + \theta_2 \sum_{j=1}^{n} W_{ij} Controls_{jt} + City_i + Year_t + \varepsilon_{it} \end{aligned} \tag{5}$$

Using the shadow banking scale of non-financial enterprises at the prefecture-level as a benchmark, this section divides the sample

3 The *New Asset Management Regulations* is short for the "Guiding Opinions on Regulating the Asset Management Business of Financial Institutions" (Yinfa [2018] No. 106) was issued by the People's Bank of China, the China Securities Regulatory Commission, the State Administration of Foreign Exchange, and the China Banking and Insurance Regulatory Commission on 27 April 2018. (please refer to http://www.pbc.gov.cn/goutongjiaoliu/113456/113469/3529600/index.html for more details)

TABLE 10 Heterogeneity Analysis: Economic Development (Hu Line).

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Distance matrix (*W*1)		Adjacency matrix (*W*2)		Economic matrix (*W*3)	
	Northwest	Southeast	Northwest	Southeast	Northwest	Southeast
SL_sum	−0.169*	−0.027**	−0.147	−0.021*	−0.162	−0.015
	(-1.78)	(-2.14)	(-1.46)	(-1.74)	(-1.58)	(-1.25)
WSL_sum	−1.095***	−0.672***	−0.297**	−0.049**	−0.192	−0.069*
	(-3.56)	(-4.34)	(-2.35)	(-2.05)	(-1.26)	(-1.77)
PopDen	13.518***	0.635	14.268***	0.497	15.162***	0.709*
	(4.44)	(1.41)	(4.30)	(1.15)	(4.49)	(1.67)
PerGDP	13.010**	1.787**	18.696***	2.233***	21.914***	0.690
	(2.50)	(2.39)	(3.52)	(3.22)	(3.77)	(1.09)
Edu	−20.831***	0.770	−31.496***	0.036	−27.102***	−0.189
	(-2.62)	(1.15)	(-3.61)	(0.06)	(-3.13)	(-0.30)
FD	15.826***	−1.119**	17.507***	−1.715***	16.779***	−2.211***
	(2.83)	(-1.99)	(2.93)	(-3.27)	(2.74)	(-4.20)
Gov	−24.009*	0.916	−24.732*	0.667	−17.654	1.792
	(-1.93)	(0.58)	(-1.85)	(0.46)	(-1.33)	(1.27)
Service	−0.272*	−0.015	−0.171	−0.020	−0.218	−0.036*
	(-1.93)	(-0.66)	(-1.17)	(-0.97)	(-1.40)	(-1.79)
ρ	−0.359***	0.895***	−0.180***	0.505***	−0.135**	0.610***
	(-3.91)	(29.06)	(-3.34)	(24.49)	(-2.39)	(23.00)
σ^2	94.646***	13.927***	105.637***	11.983***	110.374***	12.285***
	(10.95)	(34.80)	(10.93)	(33.99)	(10.97)	(34.41)
WControls	Yes	Yes	Yes	Yes	Yes	Yes
Direct effect	−0.120	−0.065***	−0.119	−0.030**	−0.151	−0.022
	(-1.24)	(-2.71)	(-1.16)	(-2.24)	(-1.45)	(-1.64)
Indirect effect	−0.854***	−7.769*	−0.254**	−0.114***	−0.189	−0.207**
	(-3.77)	(-1.88)	(-2.52)	(-2.72)	(-1.40)	(-2.24)
Total effect	−0.974***	−7.834*	−0.373***	−0.144***	−0.340**	−0.228**
	(-3.99)	(-1.89)	(-2.86)	(-2.84)	(-2.17)	(-2.30)
City FE	Yes	Yes	Yes	Yes	Yes	Yes

(Continued on the following page)

TABLE 10 (*Continued*) Heterogeneity Analysis: Economic Development (Hu Line).

	(1)	(2)	(3)	(4)	(5)	(6)
	Distance matrix (*W*1)		Adjacency matrix (*W*2)		Economic matrix (*W*3)	
	Northwest	Southeast	Northwest	Southeast	Northwest	Southeast
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	242	2,431	242	2,431	242	2,431
R^2	0.015	0.481	0.008	0.008	0.025	0.232

This table presents the regression results analyzing the impact of economic development on shadow lending activities across different regions divided by the Hu Line. Column (1), (3) and (5) present results for the Northwest, while Column (2), (4) and (6) report results for the Southeast. Two-way fixed effects are included. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

into a treatment group (regions with shadow banking scale above the 50th percentile) and a control group (regions below the 50th percentile). This classification allows for a clearer assessment of how changes in shadow banking activity, before and after the policy implementation, influence SME mortality and provides a theoretical foundation for understanding the underlying mechanism through which shadow banking affects SME survival.

Table 12 reports the regression results from the SDM-DID model estimating the effects of the *New Asset Management Rules*. The key explanatory variable in this section is the interaction term $Treat_i \times Post_t$, indicating whether a region is in the treatment group following the policy implementation. The analysis is conducted using three types of spatial weight matrices: the distance matrix (*W*1), the adjacency matrix (*W*2), and the economic matrix (*W*3). According to the regression results, the coefficients on the policy shock variable $Treat_i \times Post_t$ are 0.0139 (*W*1), 0.0152 (*W*2), and 0.0170 (*W*3), all of which are significantly positive at the 1 percent level.

These results suggest that following the implementation of the *New Asset Management Rules*, SME mortality increased significantly in regions with larger shadow banking activity (the treatment group). As shown in the earlier sections of this study, shadow banking by non-financial enterprises plays a positive role in supporting the survival of SMEs by alleviating their financing constraints. However, the primary aim of the regulatory tightening was to reduce high-risk operations within the shadow banking system and to mitigate potential systemic financial risks. In the short term, such regulatory tightening may lead to the contraction of certain shadow banking channels, thereby reducing the scale of shadow banking and the volume of capital flowing to SMEs. As a result, SMEs that previously relied on these informal financing sources may face increased financial pressure due to diminished credit availability, ultimately contributing to a rise in their mortality.

Moreover, the impact of the policy is not limited to prefecture-level cities with larger shadow banking activity. It can also extend to neighboring regions through interregional financial linkages. The spatial spillover effect is preliminarily captured by the spatially lagged interaction term ($W_{ij}(Treat_i \times Post_t)$), whose coefficients under the three spatial weight matrices are 0.1675, 0.0545, and 0.0579, respectively. All are significantly positive at the 1 percent level. This indicates that the shock generated by the *New Asset Management Rules* in regions with a high degree of shadow banking activity gradually spread to adjacent areas through financial market interdependence. Given the high degree of connectivity in financial markets and the close flow of capital across regions, the contraction of shadow banking in the affected areas exerted a secondary influence on neighboring regions, raising financing costs and financial pressure for SMEs in those areas as well. The spatial effect decomposition results in Table 11 further reveal both the direct and indirect effects of shadow banking activity on SME mortality. Under all three spatial weight matrices, the direct effects are significantly positive at the 1 percent level, with coefficients of 0.0145 for *W*1, 0.0177 for *W*2, and 0.0177 for *W*3. This indicates that policy shocks led to increased SME mortality within regions with large-scale shadow banking activity. In addition, the indirect effects are 0.2308 for *W*1, 0.0644 for *W*2, and 0.0675 for *W*3, while the total effects are 0.2453 for *W*1, 0.0821 for *W*2, and 0.0852 for *W*3. All are significantly positive at the 1 percent level. These results provide strong evidence that the impact of the policy shock was transmitted through regional financial networks, thereby influencing the survival of SMEs in neighboring areas. As a result, the mortality rates of SMEs in these surrounding regions also increased following the implementation of the *New Asset Management Rules*.

Overall, the *New Asset Management Rules* reduced the scale of shadow banking financing by strengthening regulatory oversight of high-risk activities. In the short term, this policy exerted considerable pressure on the survival environment of SMEs in regions with a high concentration of shadow banking activity and their neighboring areas, leading to increased financing costs and heightened financial stress for affected firms. However, the long-term objective of the policy is to safeguard the stability of the financial system by curbing excessive risk-taking and preventing systemic financial risks. As the policy continues to advance, the financial environment is gradually being cleansed, creating a healthier and more stable market foundation for the sustainable development of SMEs. In addition, the policy has effectively guided SMEs to reduce their dependence on shadow banking and to transition toward more sustainable and transparent financing channels. It has also contributed to a more efficient market selection process, enabling firms with sounder financial structures and stronger adaptive capacity to survive and grow. Although the mortality rate of SMEs has increased in the short term, the surviving firms are generally more resilient and competitive, thereby promoting sustained development within a more regulated financial system. In the long run, the *New Asset Management Rules* have laid

TABLE 11 Additional tests: shadow lending and total enterprises mortality rate.

Variables	(1)	(2)	(3)
	Distance matrix (*W1*)	Adjacency matrix (*W2*)	Economic matrix (*W3*)
SL	−0.014**	−0.013**	−0.014**
	(-2.00)	(-1.99)	(-2.13)
WSL	−0.271***	0.015	−0.009
	(-3.30)	(1.25)	(-0.71)
PopDen	1.203***	1.303***	0.957***
	(5.00)	(5.61)	(4.12)
PerGDP	−0.296	−0.412	0.062
	(-0.75)	(-1.11)	(0.15)
Edu	0.646*	0.703**	0.464
	(1.72)	(2.01)	(1.29)
FD	0.001	0.122	0.150
	(0.00)	(0.42)	(0.50)
Gov	−0.447	−1.297	−0.277
	(-0.52)	(-1.62)	(-0.34)
Service	−0.014	−0.010	−0.018
	(-1.22)	(-0.92)	(-1.59)
ρ	0.893***	0.467***	0.402***
	(28.79)	(23.84)	(18.40)
σ^2	4.681***	4.136***	4.427***
	(36.37)	(35.78)	(36.00)
WControls	Yes	Yes	Yes
Direct effect	−0.028***	−0.012	−0.016**
	(-2.66)	(-1.62)	(-2.26)
Indirect effect	−3.147*	0.012	−0.027
	(-1.85)	(0.64)	(-1.40)
Total effect	−3.175*	0.001	−0.043*
	(-1.86)	(0.03)	(-1.88)
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs	2,430	2,430	2,430
R^2	0.015	0.002	0.017

This table reports the results of regressions including alternative proxies for the dependent variable under different spatial weighting matrices: distance matrix, adjacency matrix, and economic matrix. The dependent variable is replaced by *MR*, refers to total enterprises mortality ratedeat across prefecture-level. Two-way fixed effects are included. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

10.3389/fphy.2025.1612653

TABLE 12 Further analysis: DID Analysis: The Impact of the "*New Asset Management Regulations*".

Variables	(1)	(2)	(3)
	Distance matrix (*W*1)	Adjacency matrix (*W*2)	Economic matrix (*W*3)
Treat × *Post*	0.0139***	0.0152***	0.0170***
	(2.80)	(3.12)	(3.53)
W(*Treat* × *Post*)	0.1675***	0.0545***	0.0579***
	(4.13)	(6.76)	(5.70)
PopDen	0.0283***	0.0286***	0.0279***
	(4.40)	(4.33)	(4.36)
PerGDP	0.0528***	0.0576***	0.0513***
	(4.96)	(5.35)	(5.29)
Edu	−0.0282***	−0.0278***	−0.0322***
	(-2.73)	(-2.72)	(-3.17)
FD	0.0383**	0.0705***	0.0451***
	(2.28)	(4.20)	(2.74)
Gov	0.0053	−0.0195	0.0016
	(0.22)	(-0.84)	(0.07)
ρ	0.2577***	0.1706***	0.1184***
	(3.07)	(8.85)	(5.21)
σ^2	0.0034***	0.0034***	0.0034***
	(36.65)	(36.53)	(36.55)
WControls	Yes	Yes	Yes
Direct effect	0.0145***	0.0177***	0.0177***
	(2.94)	(3.73)	(3.70)
Indirect effect	0.2308***	0.0644***	0.0675***
	(3.78)	(7.08)	(5.89)
Total effect	0.2453***	0.0821***	0.0852***
	(4.14)	(9.26)	(7.87)
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs	2,673	2,673	2,673
R^2	0.392	0.026	0.183

This table presents the regression results using a DID approach to examine the effect of the *New Asset Management Regulations*, which aimed to reduce systemic financial risks by regulating shadow banking activities. Cities are divided into a treatment group (with shadow banking scales above the 50th percentile) and a control group (below the 50th percentile). The policy is expected to have a more significant impact on cities with larger shadow banking activities. Two-way fixed effects are included. t-statistics are shown in parentheses. ***, **, and *indicate significance at the 1%, 5%, and 10% levels, respectively.

an institutional foundation for the healthy growth of SMEs and the high-quality development of the sector as a whole.

7 Conclusion

This study empirically examines the critical role of non-financial enterprises' shadow lending in shaping small and micro enterprises' (SMEs) survival across Chinese regions by adopting the Spatial Durbin Model (SDM). Our findings show that shadow lending by non-financial enterprises significantly reduces local SME mortality, while also generating substantial spatial spillover effects, where credit diffusion through shadow banking activities improves SME survival in neighboring regions. This demonstrates that shadow banking not only supports SMEs in their originating regions but also creates favorable externalities across interconnected regional economies.

Our findings provide significant implications for both academic inquiry and policy design. By revealing that shadow banking significantly lowers SME mortality both locally and through regional spillovers, we provide empirical support to existing theoretical perspectives on the financial intermediation role of non-financial enterprises. Moreover, our results expand upon prior studies that have focused on the spatial distribution of shadow banking activities [17, 18] by demonstrating how such geographic patterns translate into measurable impacts on SME resilience. Our analysis further complements macro-level studies (e.g., [19, 20]) that highlight the role of regulatory arbitrage and systemic risk in shadow banking development. Integrating these insights, our research bridges micro-level firm dynamics with macro-level financial policy debates, offering a geographically grounded understanding of informal finance. These findings underscore the necessity for spatially differentiated regulatory strategies that reflect the uneven role and risks of shadow banking across regions.

In addition, the heterogeneity analysis reveals that the effects of shadow lending by non-financial enterprises are more pronounced in economically developed southeastern regions, with relatively weaker effects observed in the less developed northwest. This highlights the role of regional economic development in shaping the effectiveness of shadow banking as a financing tool for SMEs. Furthermore, robustness checks confirm the persistence of these positive effects over time, with shadow banking continuing to alleviate financing pressures for SMEs over two periods.

However, the informal nature of shadow banking introduces significant systemic vulnerabilities. The unregulated flow of credit can generate financial fragility, particularly in regions where SMEs heavily depend on informal financing channels. As these credit flows deepen, financial stress in one region can spread to others through inter-regional capital linkages, heightening the risk of financial contagion.

Based on these findings, we recommend that financial regulators develop a differentiated regulatory approach. While stricter regulations are needed in economically developed regions to manage the potential risks arising from shadow banking, less developed regions may benefit from more flexible policies that support SME financing. Moreover, cross-regional financial regulatory coordination is essential to manage systemic risks and prevent the propagation of financial stress between regions. Strengthening regional risk monitoring and establishing early warning systems are critical for mitigating the broader effects of financial contagion.

Overall, this study contributes to the understanding of shadow banking's dual role in supporting SME survival while also exacerbating systemic risks. It highlights the need for tailored regulatory frameworks that consider both regional economic disparities and the cross-regional transmission of financial risks [32, 33].

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.cnrds.com/Home/Login.

Author contributions

CW: Software, Writing – original draft, Methodology. WZ: Writing – review and editing, Supervision. YM: Conceptualization, Writing – review and editing, Project administration. XX: Funding acquisition, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is supported by the National Natural Science Foundation of China (72141304, 72201190), and Key Program of the Ministry of Science and Technology, China (2022YFC3303304).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that Gen AI was used in the creation of this manuscript. During the preparation of this work, the authors used ChatGPT to assist in ONLY language refinement. The authors take full responsibility for the content of this publication.

Correction notice

This article has been corrected with minor changes. These changes do not impact the scientific content of the article.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Lin A, Peng Y, Wu X. Digital finance and investment of micro and small enterprises: evidence from China. *China Econ Rev* (2022) 75:101846. doi:10.1016/j.chieco.2022.101846

2. Carpenter RE, Petersen BC. Capital market imperfections, high-tech investment, and new equity financing. *The Econ J* (2002) 112(477):F54–F72. doi:10.1111/1468-0297.00683

3. Beck T, Demirgüç-Kunt ASLI, Maksimovic V. Financial and legal constraints to growth: does firm size matter? *The J Finance* (2005) 60(1):137–77. doi:10.1111/j.1540-6261.2005.00727.x

4. Degryse H, Matthews K, Zhao T. SMEs and access to bank credit: evidence on the regional propagation of the financial crisis in the UK. *J Financial Stab* (2018) 38:53–70. doi:10.1016/j.jfs.2018.08.006

5. Kapitsinis N, Munday M, Roberts A. Exploring a low SME equity equilibrium in Wales. *Eur Plann Stud* (2021) 29:1777–97. doi:10.1080/09654313.2021.1882945

6. Jin S, Wang W, Zhang Z. The real effects of implicit government guarantee: evidence from Chinese state-owned enterprise defaults. *Management Sci* (2023) 69(6):3650–74. doi:10.1287/mnsc.2022.4483

7. Lu Y, Guo H, Kao EH, Fung HG. Shadow banking and firm financing in China. *Int Rev Econ & Finance* (2015) 36:40–53. doi:10.1016/j.iref.2014.11.006

8. Du J, Li C, Wang Y. Shadow banking of non-financial firms: arbitrage between formal and informal credit markets in China. *J financial intermediation* (2023) 55:101032. doi:10.1016/j.jfi.2023.101032

9. Pozsar Z, Adrian T, Ashcraft A, Boesky H. Shadow banking. *Econ Policy Rev* (2013) 19(2):1–16. Available online at: www.newyorkfed.org/research/staff_reports/sr458.html.

10. Wang G-J, Feng Y, Xiao Y, Zhu Y, Xie C. Connectedness and systemic risk of the banking industry along the Belt and Road. *J Management Sci Eng* (2022) 7(3):303–29. doi:10.1016/j.jmse.2021.12.002

11. Bester H. The role of collateral in credit markets with imperfect information. *Eur Econ Rev* (1987) 31(4):887–99. doi:10.1016/0014-2921(87)90005-5

12. Agostino M, Trivieri F. Does trade credit play a signalling role? Some evidence from SMEs microdata. *Small Business Econ* (2014) 42:131–51. doi:10.1007/s11187-013-9478-8

13. Andrieu G, Staglianò R, Van Der Zwan P. Bank debt and trade credit for SMEs in Europe: firm-, industry-, and country-level determinants. *Small Business Econ* (2018) 51:245–64. doi:10.1007/s11187-017-9926-y

14. Li S, Wang C. Network structure, portfolio diversification and systemic risk. *J Management Sci Eng* (2021) 6(2):235–45. doi:10.1016/j.jmse.2021.06.006

15. Berger AN, Udell GF. The economics of small business finance: the roles of private equity and debt markets in the financial growth cycle. *J banking & Finance* (1998) 22(6-8):1–69. doi:10.17016/feds.1998.15

16. Wang J, Liu G, Xu X, Xing X. Credit risk prediction for small and medium enterprises utilizing adjacent enterprise data and a relational graph attention network. *J Management Sci Eng* (2024) 9(1):177–92. doi:10.1016/j.jmse.2023.11.005

17. Yuan C, Jiang H, Ren J, Kapitsinis N. Collateral monetary policy, regional financial development, and nonfinancial firms' shadow banking activities. *China & World Economy* (2024) 32:166–96. doi:10.1111/cwe.12552

18. Zhang Q, Que J, Qin X. Regional financial technology and shadow banking activities of non-financial firms: evidence from China. *J Asian Econ* (2023) 86:101606. Article No. 101606. doi:10.1016/j.asieco.2023.101606

19. Buchak G, Matvos G, Piskorski T, Seru A. Fintech, regulatory arbitrage, and the rise of shadow banks. *J Financial Econ* (2018) 130(3):453–83. doi:10.1016/j.jfineco.2018.03.011

20. Yuan C, Zhu Y, Kapitsinis N. Collateral monetary policy, shadow banking and bank risk: evidence from China. *China Finance Rev Int* (2025). doi:10.1108/cfri-07-2024-0423

21. Ayyagari M, Demirgüç-Kunt A, Maksimovic V. Formal versus informal finance: evidence from China. *The Rev Financial Stud* (2010) 23(8):3048–97. doi:10.1093/rfs/hhq030

22. Petersen MA, Rajan RG. The benefits of lending relationships: evidence from small business data. *The J Finance* (1994) 49(1):3–37. doi:10.1111/j.1540-6261.1994.tb04418.x

23. Rice T, Strahan PE. Does credit competition affect small-firm finance? *The J Finance* (2010) 65(3):861–89. doi:10.1111/j.1540-6261.2010.01555.x

24. Porter ME. Location, competition, and economic development: local clusters in a global economy. *Econ Dev Q* (2000) 14(1):15–34. doi:10.1177/089124240001400105

25. Anselin L. *Spatial econometrics: methods and models*. Dordrecht: Kluwer Academic Publishers (1988).

26. Fritsch M, Mueller P. Effects of new business formation on regional development over time. *Reg Stud* (2004) 38(8):961–75. doi:10.1080/0034340042000280965

27. Allen F, Qian Y, Tu G, Yu F. Entrusted loans: a close look at China's shadow banking system. *J Financial Econ* (2019) 133(1):18–41. doi:10.1016/j.jfineco.2019.01.006

28. LeSage JP, Pace RK. Spatial econometric models. In: MM Fischer, A Getis, editors. *Handbook of applied spatial analysis: Software tools, methods and applications*. Berlin, Heidelberg: Springer (2009). p. 355–76.

29. Elhorst JP. Applied spatial econometrics: raising the bar. *Spat Econ Anal* (2010) 5(1):9–28. doi:10.1080/17421770903541772

30. Moran PA. Notes on continuous stochastic phenomena. *Biometrika* (1950) 37(1/2):17–23. doi:10.2307/2332142

31. Anselin L. Local indicators of spatial association—LISA. *Geographical Anal* (1995) 27(2):93–115. doi:10.1111/j.1538-4632.1995.tb00338.x

32. Yang L, Hong Y, Zhang X, Zhang Q. Can FinTech alleviate the long-term use of short-term debts? Evidence from China. *China Finance Rev Int* (2024). doi:10.1108/cfri-06-2024-0329

33. Jiang K, Xie X, Xiao Y, Ashraf BN. The value of corporate digital transformation: evidence from bond pricing. *China Finance Rev Int* (2024) 15(1):43–66. doi:10.1108/cfri-05-2024-0241

34. Hu H. The distribution of China's population. *Acta Geographica Sinica* (1935) 2(2):33–74.

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 03 June 2025
DOI 10.3389/fphy.2025.1612695

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Qinnan Jiang,
Changsha University of Science and Technology, China
Mingyu Shu,
Anhui University, China

*CORRESPONDENCE
An Li,
lianacd@126.com

RECEIVED 16 April 2025
ACCEPTED 20 May 2025
PUBLISHED 03 June 2025

CITATION
Li J, Li Y, Lai F and Li A (2025) Global geopolitical risk and industrial tail risk in the Chinese stock market: a quantile-on-quantile connectedness approach.
Front. Phys. 13:1612695.
doi: 10.3389/fphy.2025.1612695

COPYRIGHT
© 2025 Li, Li, Lai and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Global geopolitical risk and industrial tail risk in the Chinese stock market: a quantile-on-quantile connectedness approach

Jing Li[1], Yunzhong Li[2], Fujun Lai[3] and An Li[4]*

[1]Business School, Xiangtan University, Xiangtan, China, [2]National Engineering Laboratory for Big Data Analysis and Application, School of Mathematical Sciences, Peking University, Beijing, China, [3]School of Finance, Yunnan University of Finance and Economics, Kunming, China, [4]Chinese Academy of Finance and Development, Central University of Finance and Economics, Beijing, China

Global geopolitical risk (GPR) has increasingly become a pivotal driver of financial market volatility, understanding the impact of GPR on market tail risk is crucial, particularly as traditional models often overlook the complex, nonlinear dynamics exacerbated by geopolitical shocks. This study offers an in-depth examination of the quantile-dependent spillover connectedness between GPR and the tail risk of 18 industries in the Chinese stock market. By using a quantile-on-quantile (QQ) connectedness approach, we investigate how shocks at varying quantiles propagate through the system, thereby uncovering nonlinear dynamics often obscured by traditional mean-variance models. Our findings reveal a distinct "U-shaped" quantile dependence, where extreme quantiles (5% and 95%) exhibit significantly heightened sensitivity to GPR compared to mid-range quantiles. Additionally, a net directional analysis demonstrates that industries with global integration or resource intensity (such as Manufacturing, Mining, and IT) typically serve as net risk receivers during geopolitical turbulence, while certain sectors (notably Finance) may act as net risk senders under specific conditions. A dynamic connectedness analysis further indicates that pivotal geopolitical events, including the 2018 China-U.S. trade war, the COVID-19 pandemic and the 2022 Russia-Ukraine conflict, act as junctures that intensify tail risk transmission. Collectively, these insights emphasize the necessity of quantile-specific risk monitoring and underscore the value of tailored policy interventions to mitigate severe downside risks amid escalating global uncertainties.

KEYWORDS

global geopolitical risk, tail risk, quantile-on-quantile connectedness, Chinese stock market, extreme market conditions

1 Introduction

Global geopolitical risk (GPR) has emerged as a pivotal force driving financial market volatility, encompassing events such as wars, significant political tensions, economic sanctions, and trade barriers [1, 2]. Such disruptions profoundly influence investor sentiment and capital flows, triggering cross-market spillovers and amplifying uncertainty [3]. Amid escalating military conflicts, prolonged trade disputes, and pandemic-induced

shocks, policy uncertainty has surged to unprecedented levels, heightening systemic vulnerabilities across both developed and emerging equity markets [4]. Consequently, the World Economic Forum has identified GPR as a foremost global threat, advocating for strengthened market oversight to curb downside risks [1].

Concurrently, tail risk—defined as the likelihood of extreme losses stemming from rare, high-impact events—has gained increasing prominence in risk management and academic research [5, 6]. Traditional financial models, based on mean-variance frameworks, often fail to capture the heavy-tailed, nonlinear patterns that intensify during disruptive geopolitical events [7, 8]. When "black swan" or "gray rhino" events—such as wars or sweeping sanctions—unfold, stock returns exhibit heightened asymmetry, fostering widespread spillovers that exacerbate systemic risk [1]. This dynamic is particularly salient in emerging markets like China, where retail investors' herding behavior tends to amplify extreme selloffs amid informational uncertainty [9].

Despite growing scholarly attention to GPR [2, 10], most studies focus on average returns or aggregate volatility, overlooking the quantile-specific mechanisms through which GPR amplifies tail risk. While some researchers note that cross-industry and cross-regional linkages surge during severe market downturns, evidence on the underlying non-linear pathways remains limited [11, 12]. This gap prompts a critical question: how can we effectively quantify the impact of global geopolitical shocks on the lower tails of the return distribution? Addressing this is vital for policymakers seeking to safeguard systemic stability amid rising global tensions.

To bridge this gap, we employ a quantile-on-quantile (QQ) connectedness approach to explore spillovers between GPR and the tail risks of 18 Chinese stock market industries. The QQ connectedness approach systematically traces how shocks at specific GPR quantiles transmit to corresponding industrial tail risk quantiles [13–15], thereby revealing nonlinear dynamics and asymmetries in risk transmission. By accounting for both lower- and upper-tail conditions, it reveals nuanced patterns and U-shaped dependencies often overlooked by simpler linear or single-quantile methodology [16]. Notably, our approach highlights key scenarios—such as the 2018 China–U.S. trade war, COVID-19 pandemic, and the 2022 Russia–Ukraine conflict—where geopolitical shocks disproportionately affect the left tails of China's stock market [14, 17].

Using monthly data from 2010 to 2023, we derive three principal findings. First, a pronounced "U-shaped" dependence emerges between GPR and Chinese industrial tail risk, with extreme quantiles (5% and 95%) exhibiting markedly stronger connectedness than mid-quantiles. Second, net directional analysis reveals that globally integrated, resource-intensive sectors—such as Manufacturing, Mining, and IT—consistently act as net risk receivers under geopolitical stress, while Finance occasionally emerges as a net risk sender, amplifying shocks under certain conditions. Third, dynamic connectedness peaks during major geopolitical events, confirming that incidents like the 2018 trade war, the COVID-19 pandemic, and the 2022 Russia–Ukraine conflict serve as pivotal moments that heighten tail risk transmission. These findings highlight the critical need for quantile-specific risk monitoring and advocate for targeted policy measures to address heightened geopolitical risk periods, aligning with calls for macroprudential strategies to manage financial turbulence [18].

Our study offers three major contributions. First, it broadens the scope of research on GPR by exploring its dynamic interplay with tail risks across 18 diverse industries in the Chinese stock market, uncovering how the connectedness evolves across different quantiles and over time. This enriches the understanding of how geopolitical uncertainties shape extreme risk profiles in a major emerging market. Second, while prior studies [18–23] often limit their analysis to spillovers within uniform quantiles, overlooking cross-quantile dynamics, this paper employs the cutting-edge QQ connectedness approach to provide a detailed and precise depiction of risk transmission between GPR and industrial tail risks under varying market conditions. Third, unlike earlier research that typically examines broad market indices or isolated sectors, we develop an innovative framework that categorizes 18 industries into dominant, basic support, and modern service groups, enabling a granular analysis of how geopolitical shocks differentially propagate across distinct economic segments.

The paper is structured as follows: Section 2 reviews the theoretical background and literature. Section 3 details the methodology and data. Section 4 presents empirical findings on tail risk connectedness under geopolitical shocks. Section 5 concludes with policy implications and future research directions.

2 Theoretical background and literature review

2.1 Theoretical background

The global financial landscape has grown increasingly volatile due to escalating geopolitical tensions, including wars, terrorist attacks, and diplomatic crises [10, 24]. GPR, characterized as the volatility and uncertainty stemming from such events, profoundly influences policy frameworks and economic stability, thereby reshaping the risk profiles of financial markets worldwide. Within this context, tail risk—defined as the probability of extreme negative returns at the distribution's far end—emerges as a critical concern, particularly susceptible to GPR's disruptive effects.

Theoretically, GPR impacts tail risk through three primary channels: uncertainty, risk aversion, and spillover or contagion. First, the uncertainty channel suggests that GPR amplifies ambiguity in policy and macroeconomic environments, elevating risk premia. Pástor and Veronesi [24] argue that heightened policy uncertainty increases expected returns, while Baker et al. [10] find that the Economic Policy Uncertainty (EPU) Index is positively associated with stock market volatility and reduced investment. Geopolitical events, by intensifying uncertainty, may prompt investors to demand higher risk premia or adopt a cautious 'wait-and-see' stance, heightening the potential for tail events [25].

Second, the risk aversion channel posits that rising GPR diminishes investors' willingness to bear risk. During periods of geopolitical upheaval, investors often shift from equities to safe-haven assets like gold [26], triggering a flight-to-safety effect. This shift results in equity sell-offs, reduced liquidity, and amplified price declines, thereby increasing the likelihood of extreme tail losses. Intriguingly, in markets like China, GPR can attract capital inflows, reducing firms' financing costs and spurring investment

[27, 28]. These results challenge conventional 'risk-off' assumptions and underscore GPR's multifaceted influence.

Third, the spillover and contagion channel highlights GPR's capacity to propagate systemic risk, particularly during large-scale events such as wars. Lai et al. [29] emphasize that financial market interconnectedness exacerbates risk contagion, with geopolitical shocks intensifying tail-risk spillovers across global markets [30, 31]. Collectively, these mechanisms suggest that GPR skews the risk distribution of financial markets, rendering extreme adverse outcomes more probable, especially in emerging economies like China, where market dynamics and investor behavior amplify such effects.

2.2 Literature review

Extensive research has explored the relationship between GPR and financial market dynamics, focusing on uncertainty, market responses, and firm-level impacts. This body of work can be categorized into macro-level analyses of market risk and micro-level studies of firm behavior, offering a foundation for understanding GPR's role in tail risk escalation.

2.2.1 Uncertainty and market outcomes

Early research has established that policy uncertainty significantly shapes financial markets by influencing asset pricing, economic behavior, and investor strategies. Pástor and Veronesi [24] demonstrate that policy uncertainty elevates required risk premia, driving up expected returns. Building on this, Baker et al. [10] deepen the analysis by linking uncertainty to greater stock market volatility and economic downturns, highlighting how it amplifies risk across financial systems. At the firm level, Gulen and Ion [32] found that companies respond by scaling back investment, a reaction that deepens economic slowdowns and further pressures asset prices. Meanwhile, in derivative markets, Kelly et al. [33] observed that investors, facing political upheavals, turn to options for tail risk insurance, reflecting efforts to shield themselves from uncertainty's extreme outcomes. Though these studies primarily address policy uncertainty rather than GPR specifically, they underscore uncertainty's pivotal role in shaping financial market behavior during crises.

2.2.2 Global geopolitical risk research

The influence of GPR on financial markets has gained significant attention in recent research, as global tensions increasingly shape economic stability and investor behavior. A key development in this field is the GPR Index, introduced by Caldara and Iacoviello [2], which offers a precise tool to quantify geopolitical tensions and supports deeper empirical analysis. Research shows that GPR has a profound impact, often exceeding that of traditional economic risks. For instance, Triki and Ben Maatoug [26] highlight how elevated GPR triggers a flight-to-safety effect, driving investors away from equities toward safe-haven assets like gold, which depresses stock prices and heightens market instability. This effect is compounded by findings from Balcilar et al. [34], who demonstrate that GPR not only increases stock market volatility but also extends its duration, adding a persistent layer of uncertainty that conventional risk models struggle to capture. Demirer et al. [35] also emphasize GPR's broader role in shaping global stock market dynamics and influencing asset pricing, pointing to its systemic importance across markets. Together, these insights reveal that GPR is not just a temporary disruption but a fundamental force that reshapes financial ecosystems, with lasting effects on asset classes and market behavior worldwide.

2.2.3 Evidence from China

In emerging markets like China, GPR interacts with local dynamics to shape financial outcomes in complex ways. Research shows that GPR significantly increases the downside tail risk of Chinese equities, raising the likelihood of extreme losses during periods of heightened tension [36]. This elevated risk environment may explain the 'low-exposure premium' identified by Zhang et al. [3], where firms with limited GPR exposure achieve higher returns amid geopolitical stress. Additionally, Pan et al. [4] find that GPR amplifies inter-sectoral volatility spillovers, highlighting the nonlinear and interconnected nature of market responses. Together, these studies underscore GPR's multifaceted influence on Chinese markets, offering critical insights for risk management.

2.2.4 Tail-risk measurement

Traditional systemic risk metrics, such as CoVaR [5] and the spillover index [37, 38], typically focus on mean-variance relationships, overlooking tail dependencies. Recent studies leverage quantile-based methods to address this gap. Zhao et al. [31] show that inter-market contagion varies significantly between normal and extreme conditions, validating the use of quantile approaches for extreme risk analysis. Compared to standard vector autoregression, the QVAR-based framework excels at capturing asymmetric and dynamic tail transmissions, particularly during large-scale disruptions [18, 19]. These developments enhance our understanding of GPR's impact on extreme financial outcomes [13].

In summary, the literature confirms that GPR exacerbates tail risk by amplifying uncertainty, curbing risk tolerance, and intensifying contagion effects. Yet, a comprehensive quantile-based exploration of China's equity market remains scarce, particularly regarding variations in extreme risk transmission. While CoVaR and mean-variance frameworks provide value, they fall short in elucidating the specific mechanisms of tail events triggered by geopolitical shocks. Addressing this gap, our study employs a QQ connectedness approach to precisely delineate how GPR influences the tail-risk distribution of Chinese stocks, enriching theoretical insights and offering practical guidance for mitigating downside risks.

3 Methodology and data

3.1 Methodology

Building upon the quantile connectedness framework by Chatziantoniou et al. [39] and extending the quantile-on-quantile advancement introduced by Gabauer and Stenfors [13], this study employs a QQ connectedness approach to investigate the dynamic interplay between GPR and the tail risks of 18 industries in China's stock market. Such a model is particularly well-suited for capturing

TABLE 1 Summary statistics.

Risks	Mean	Variance	Skewness	Ex.Kurtosis	JB	ERS	Q (20)	Q2 (20)
GPR	4.567	0.06	1.277	3.245	119.371***	−3.310***	177.322***	171.473***
AFAHF	0.312	0.003	−0.060	−1.138	9.167***	−1.375	662.695***	643.407***
MIN	0.401	0.005	−0.546	−0.758	12.382***	−1.345	783.940***	745.486***
MFG	0.361	0.001	−0.373	−0.460	5.385*	−2.197**	577.546***	577.802***
EHGW	0.364	0.003	−0.575	−0.012	9.247***	−0.884	787.932***	761.179***
CON	0.387	0.003	−0.596	−0.521	11.859***	−2.146**	498.655***	482.672***
WRT	0.356	0.002	−0.408	−0.587	7.079**	−2.044**	648.066***	657.844***
TSPS	0.397	0.003	−0.370	−0.782	8.122**	−1.478	848.859***	855.853***
A&C	0.377	0.004	−0.033	−1.247	10.911***	−2.209**	722.205***	721.914***
ITSI	0.326	0.002	−0.228	−0.803	5.980**	−2.104**	613.204***	630.915***
FIN	0.408	0.001	−0.467	−0.185	6.349**	−1.497	302.118***	295.460***
RE	0.385	0.002	−0.840	0.107	19.844***	−2.158**	652.161***	618.874***
LBS	0.343	0.002	−0.132	−0.555	2.639	−3.029***	280.138***	291.049***
SRTS	0.293	0.001	−0.116	−0.004	0.374	−3.603***	89.062***	91.488***
WEPFM	0.324	0.002	−0.360	0.090	3.694	−1.735*	489.351***	519.898***
EDU	0.314	0.004	0.663	−0.064	12.336***	−0.594	559.048***	543.886***
HSW	0.338	0.002	−0.253	−0.204	2.079	−2.866***	221.898***	221.571***
CSE	0.320	0.003	−0.208	−0.513	3.053	−2.460**	394.130***	414.666***
COM	0.379	0.003	−0.420	−0.648	7.882**	−2.431**	417.847***	412.428***

Notes: *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

quantile-specific interdependencies and nonlinear spillovers among GPR and industrial tail risks across different market conditions.

We consider a k-dimensional vector x_t of endogenous variables that includes the GPR index and tail risk measures for 18 industry sectors in the Chinese stock market (hence k = 19). At a given quantile level τ in [0,1], the QVAR(p) model can be written as Equation 1:

$$x_t = \mu(\tau) + \sum_{j=1}^{p} B_j(\tau) x_{t-j} + u_t(\tau) \tag{1}$$

where p is the lag order of the VAR, $\mu(\tau)$ is a $k \times 1$ vector of quantile-specific intercepts (conditional means), $B_j(\tau)$ are $k \times k$ coefficient matrices at lag j, and $u_t(\tau)$ is a $k \times 1$ vector of residuals at quantile τ. In our setup, all components of x_t (GPR and the 18 industrial tail risk variables) are treated as endogenous. The QVAR framework thus enables modeling of the dynamic interactions between GPR and industrial tail risks at the τ-th quantile of their joint distribution. Following the Wold decomposition theorem, the above QVAR can be transformed into an equivalent quantile vector moving average (QVMA) representation in Equation 2:

$$x_t = \mu(\tau) + \sum_{j=1}^{p} B_j(\tau) x_{t-j} + u_t(\tau) = \mu(\tau) + \sum_{i=0}^{\infty} A_i(\tau) u_{t-1}(\tau) \tag{2}$$

where $A_i(\tau)$ are $k \times k$ coefficient matrices in the moving-average form. These $A_i(\tau)$ matrices describe how a unit shock at the τ quantile propagates through the system over time. The Wold expansion thus converts the QVAR into an (infinite-order) QVMA form, which facilitates analysis of the dynamic effects of shocks across different quantile conditions and forms the basis for measuring spillover effects and connectedness among the variables.

Next, to quantify how shocks emanating from one variable affect the forecast error variance of another, the present study applies the generalized forecast error variance decomposition (GFEVD). Specifically, for a forecast horizon of F periods, the fraction of the forecast error variance in variable i attributable to shocks in variable j at quantile τ is given by Equation 3:

$$\phi^{g}_{i \leftarrow j,\tau}(F) = \frac{\sum_{f=0}^{F-1} \left(e_i' A_f(\tau) H(\tau) e_j \right)^2}{H_{ii}(\tau) \sum_{f=0}^{F-1} \left(e_i' A_f(\tau) H(\tau) A_f(\tau)' e_i \right)} \tag{3}$$

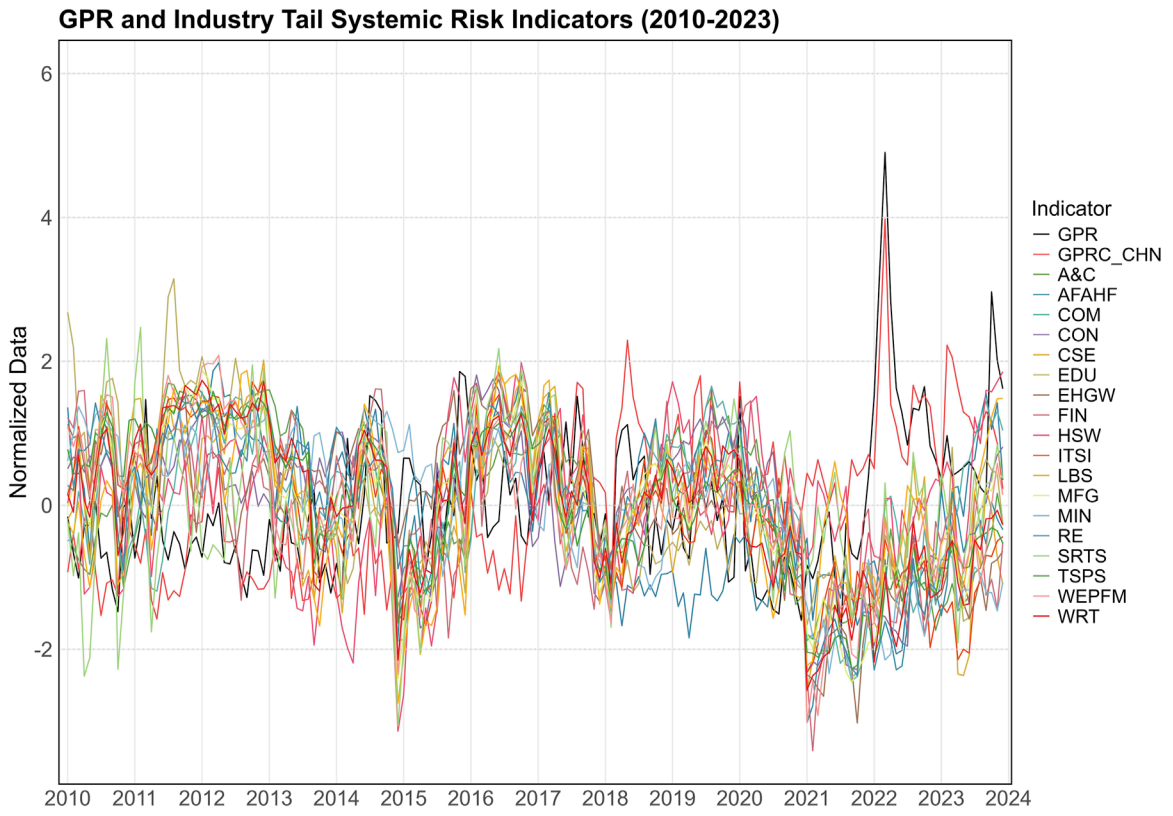


FIGURE 1
Time-varying of GPR and the eighteen industrial tail risks.

where $\phi^{g}_{i\leftarrow j,\tau}(F)$ denotes the fraction of variable i's F-period ahead forecast error variance that is attributable to shocks in variable j (at quantile level τ). Here $H(\tau) = Cov[u_t(\tau),]$ is the covariance matrix of the shock vector $u_t(\tau)$, with $H_{ii}(\tau)$ being the variance of the shock to variable i. e_i and e_j are k-dimensional selection unit vectors (with a one in the i-th or j-th position and 0 elsewhere) that pick out the respective elements. The use of a generalized FEVD [40] ensures that the variance decomposition is invariant to the ordering of variables in the system.

Based on the above variance decomposition, we proceed to construct the pairwise and system-wide connectedness measures. We first normalize the GFEVD contributions to obtain a relative influence of each shock. The normalized (or "scaled") variance share of shocks from j in explaining the variance of i is defined in Equation 4:

$$gSOT_{i\leftarrow j,\tau}(F) = \frac{\phi^{g}_{i\leftarrow j,\tau}(F)}{\sum_{j=1}^{k}\phi^{g}_{i\leftarrow j,\tau}(F)} \tag{4}$$

which represents the proportion of i's forecast error variance (at horizon F and quantile τ) contributed by shocks originating from variable j. By construction, we have $\sum_{j=1}^{k} gSOT_{i\leftarrow j,\tau}(F) = 1$ for each i, meaning the contributions of all k variables (including i itself) to the variance of i sum up to 100%. Using these normalized contributions, we then calculate the directional connectedness indices, which characterize each variable's influence on others and its vulnerability to shocks from others.

The "TO" directional connectedness of variable i measures the overall impact of shocks in i on all other variables. It is computed as Equation 5:

$$S^{gen,to}_{i\rightarrow\cdot,\tau} = \sum_{k=1,i\neq j}^{k} gSOT_{k\leftarrow i,\tau} \tag{5}$$

which sums the normalized contributions of i's shocks to the variance of every other variable j ($i \neq j$).

Conversely, the "FROM" directional connectedness of variable i captures the total influence of shocks from all other variables on i's variance, as shown in Equation 6:

$$S^{gen,from}_{i\leftarrow\cdot,\tau} = \sum_{k=1,i\neq j}^{k} gSOT_{k\rightarrow i,\tau} \tag{6}$$

which is the aggregated portion of the forecast error variance explained by shocks from all other $k-1$ variables. These two metrics $S^{gen,to}_{i\rightarrow\cdot,\tau}$ and $S^{gen,from}_{i\leftarrow\cdot,\tau}$ indicate, respectively, how much variable i contributes to volatility in the system and how susceptible i is to spillovers from the rest of the system.

Finally, we derive the Total Connectedness Index (TCI) to gauge the overall interconnectedness of the system at quantile level τ. The TCI is defined as the average of the off-diagonal spillover contributions (with an adjustment factor to account for self-contributions), given by Equation 7:

$$TCI_{\tau}(F) = \frac{k}{k-1}\sum_{k=1}^{k} S^{gen,from}_{i\leftarrow\cdot,\tau} \equiv \frac{k}{k-1}\sum_{k=1}^{k} S^{gen,to}_{i\rightarrow\cdot,\tau} \tag{7}$$

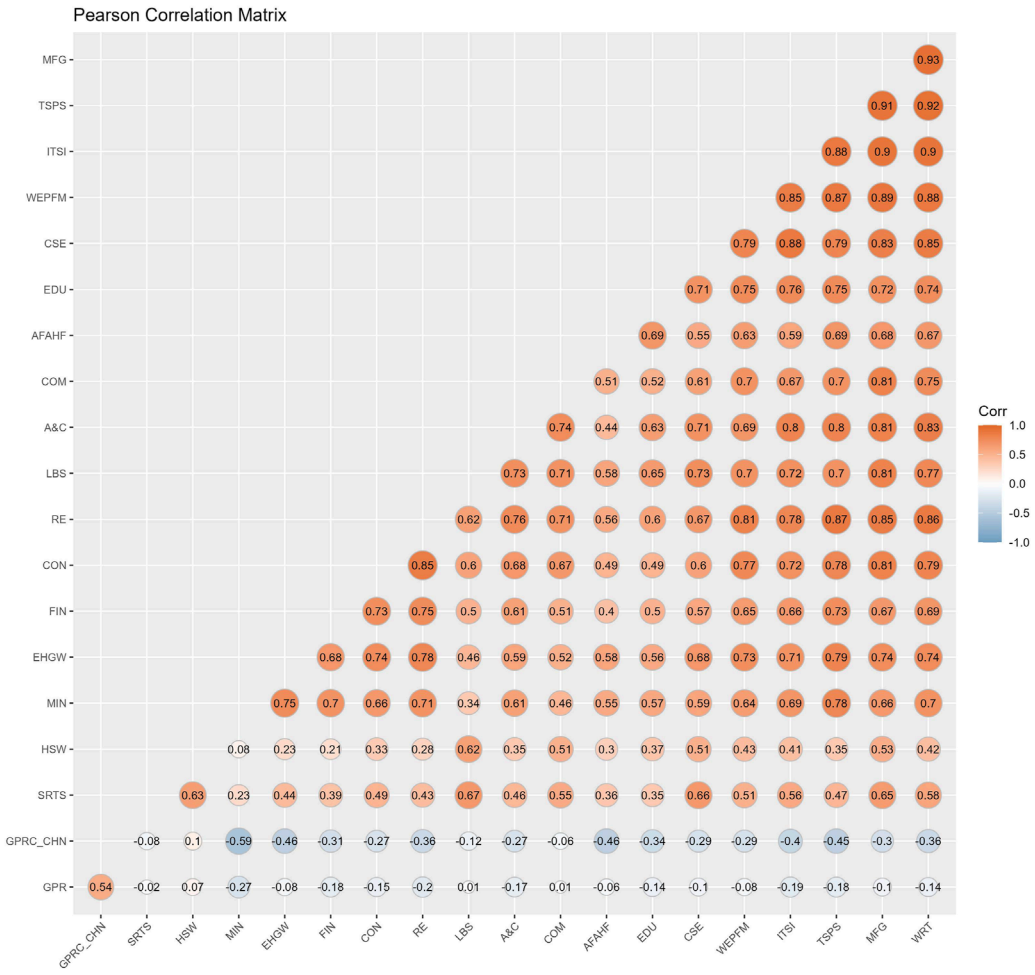


FIGURE 2
Correlation matrix between GPR and the eighteen industrial tail risks.

This adjusted TCI measures the proportion of the total forecast error variance in the system that is due to cross-variable shocks (i.e., spillovers), as opposed to idiosyncratic (own) effects. A higher TCI_τ indicates a more strongly connected system at the τ quantile. By evaluating the TCI at tail quantiles, we can capture how extreme market conditions affect the overall spillover intensity.

For the empirical implementation, we calculate the above connectedness measures within a rolling-window framework to uncover their evolution over time. In practice, we re-estimate the QVAR model on a moving sample window of fixed length, and compute the GFEVD for a forecast horizon of F periods at each window. The window size and the lag order p of the QVAR are chosen using the Bayesian Information Criterion (BIC) to ensure a balance between model fit and parsimony. This procedure generates time-varying connectedness measures, encompassing TCI and directional connectedness, that reveal how linkages between GPR and Chinese industries' tail risk evolve over time.

3.2 Data

3.2.1 Global geopolitical risk

The data comprises the GPR index and eighteen industry-specific tail systemic risk indicators as the principal variables. The GPR index, developed by Caldara and Iacoviello [2], measures geopolitical tensions by surveying the frequency of newspaper articles reporting disruptive geopolitical events, such as wars, terrorist threats, and heightened military conflict, and weighting them against the total number of news reports. Higher GPR values imply higher degrees of geopolitical tension. The sample spans 2010–2023, a period marked by multiple significant geopolitical disruptions.

3.2.2 Tail risk

To capture extreme downside linkages between each industry and the market, the study constructs eighteen industry-level tail

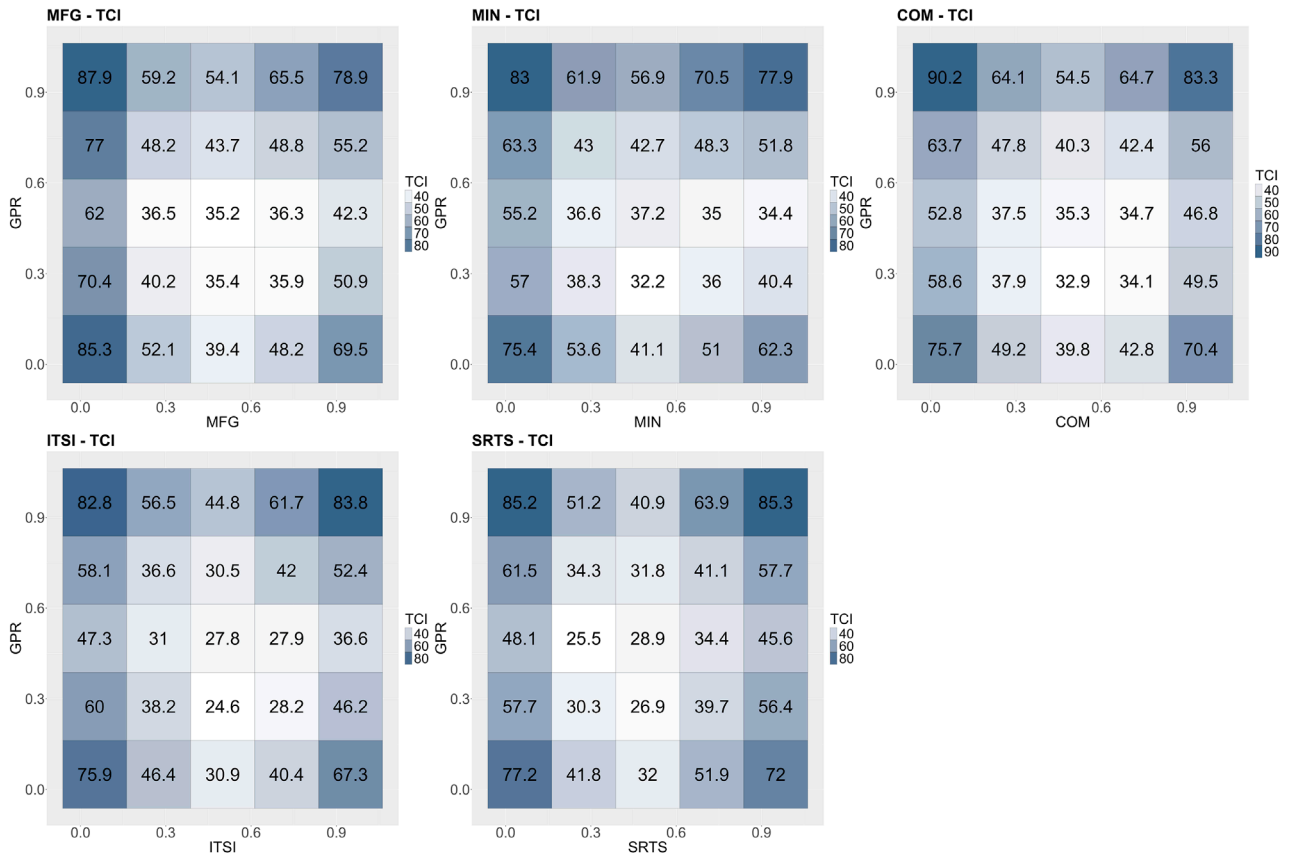


FIGURE 3
Average quantile-on-quantile connectedness between GPR and the tail risk of the dominant industries.

risk indicators. Specifically, for each of the 18 major industries in China's A-share market, the lower-tail dependence (λ_L) between that industry's returns and the broad market is estimated via the Symmetrized Joe–Clayton (SJC) copula model [41]. This model efficiently handles asymmetric dependence in the upper and lower tails, focusing on downside co-movements for systemic risk analysis. A higher λ_L implies an industry is more susceptible to market crashes, reflecting heightened systemic vulnerability. Each λ_L series is obtained using rolling windows of daily returns, permitting time variation in tail dependence. The daily stock price data are sourced from Wind and CSMAR databases, with the Shanghai Stock Exchange Composite Index serving as the market proxy. Each tail risk indicator lies between 0 and 1, where larger values signal greater systemic risk in the left tail.

3.2.3 Summary statistics

Table 1 presents key descriptive statistics for the GPR index and the 18 industrial tail risk measures from 2010 to 2023, including means, variances, skewness, kurtosis, the Jarque–Bera (JB) normality test, and the Elliott–Rothenberg–Stock (ERS) test for unit roots. On average, the GPR index stands at 4.567 with a variance of 0.060, reflecting substantial volatility relative to the tail risk variables, whose variances range from 0.001 to 0.005. Mean tail risk coefficients typically lie between 0.293 and 0.408, suggesting that roughly 30%–40% of an industry's downside losses coincide with market downturns.

Statistical tests confirm that the GPR distribution is highly non-normal, with a right skew of 1.277 and kurtosis of 3.245. By contrast, the tail risk indicators generally exhibit mild negative skewness and somewhat lower kurtosis, often indicating flatter distributions compared to the normal benchmark. Moreover, Jarque–Bera tests overwhelmingly reject normality for both GPR and most tail risk series. ERS unit root tests suggest that GPR and a majority of the industrial tail risk measures are stationary, allowing the study to proceed without marked concerns over non-stationarity.

3.2.4 Temporal patterns in GPR and industrial tail risk

Figure 1 plots the normalized time series of GPR alongside the eighteen industrial tail risks. Notably, in mid-2015, a significant domestic equity market crash in China drove sharp spikes in industry-wide tail risk, but GPR remained relatively subdued, implying a primarily domestic rather than geopolitical event. In early 2020, the COVID-19 pandemic triggered a generalized increase in tail risk and a moderate uptick in GPR. The Russia–Ukraine conflict in early 2022 propelled GPR to its highest level, again elevating industrial tail risk. These episodes demonstrate that geopolitical shocks typically coincide with concurrent increases in GPR and tail risk. In contrast, episodes driven by internal market factors

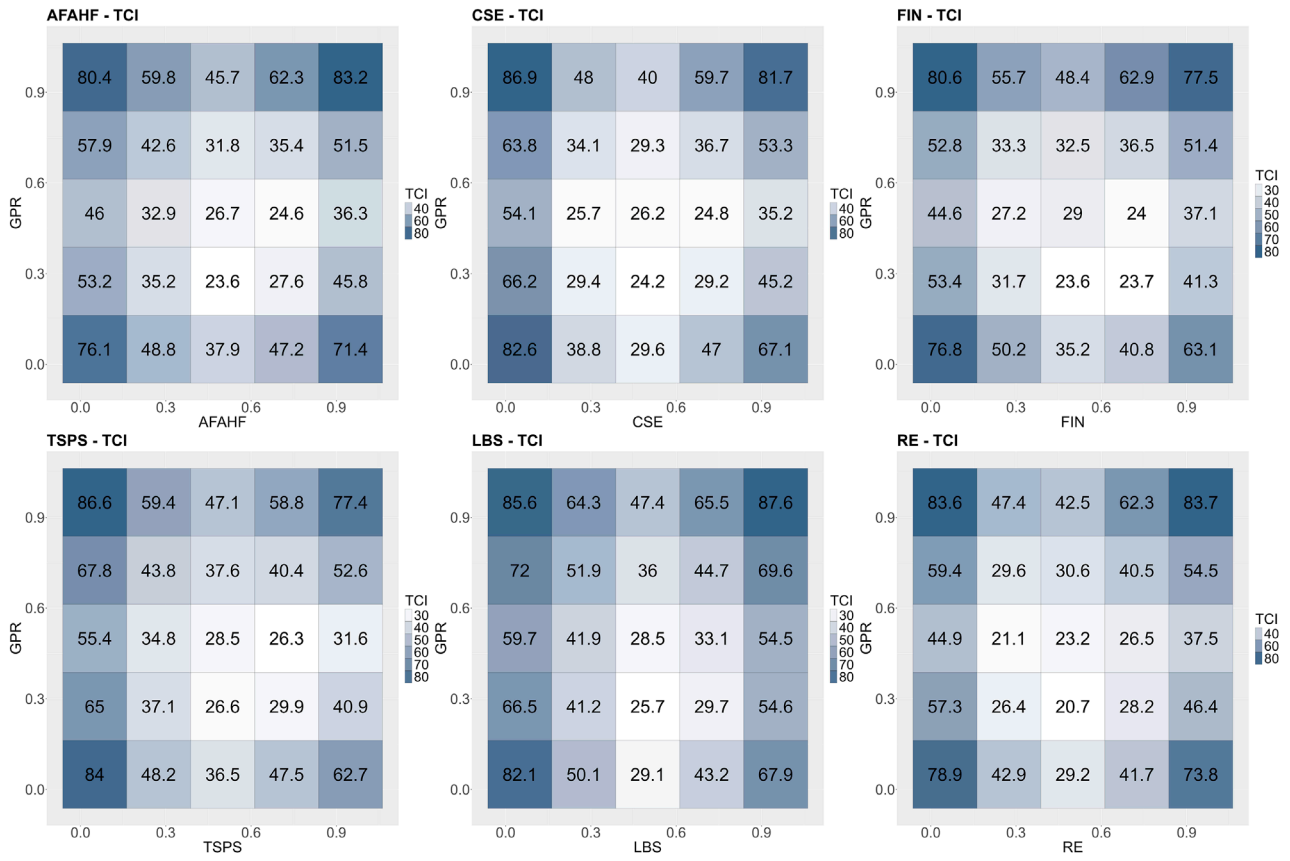


FIGURE 4
Average quantile-on-quantile connectedness between GPR and the tail risk of the basic support industries.

exert a stronger effect on tail risks without substantially altering geopolitical risk.

3.2.5 Pearson correlation

Figure 2 presents a Pearson correlation heatmap for GPR and the tail risks. Correlations among the tail risks are predominantly high and positive, approaching 1, suggesting a common driver of systemic risk across sectors. By contrast, correlations between the GPR index and each tail risk remain low, mostly near zero or mildly positive, signifying a minimal linear relationship. This low correlation underscores the necessity of employing the QQ connectedness approach to detect non-linear and asymmetric linkages, especially under severe market or geopolitical conditions.

4 Empirical analysis

This section presents the empirical findings on the interconnectedness between GPR and the tail risk of 18 industries within the Chinese stock market across various quantiles. First, we examine the Quantile-on-Quantile Total Connectedness (TCI) to evaluate the connectedness of GPR and tail risk across all industries under different quantiles. Second, we employ the Net Quantile-on-Quantile Total Directional Connectedness (NET) to identify whether each industry acts as a net risk sender or receiver. Third, we focus on the Dynamic Correlation Total Connectedness (DYNAMIC) to highlight how these interdependencies evolve in response to major geopolitical events. Robustness checks are discussed at the end.

In this study, the 18 industries are categorized into three distinct groups to reveal the paths of risk transmission and potential tail interlinkages under geopolitical shocks. This classification is based on each industry's functional positioning, upstream and downstream linkages, and overall contribution to the economy, aligning with official standards such as the Guidelines for Industry Classification of Listed Companies (2012 Revision) issued by the China Securities Regulatory Commission and the three-sector classification framework in GB/T 4754-2017 by the National Bureau of Statistics. Specifically, the dominant industries encompass Manufacturing (MFG), Mining (MIN), Comprehensive (COM), Information Transmission, Software, and Information Technology Services (ITSI), and Scientific Research and Technical Services (SRTS). The basic support industries include Agriculture, Forestry, Animal Husbandry, and Fishery (AFAHF), Culture, Sports, and Entertainment (CSE), Finance (FIN), Transportation, Storage, and Postal Services (TSPS), Leasing and Business Services (LBS), and Real Estate (RE). Modern service industries comprise Electricity, Heat, Gas, and Water Production and Supply (EHGW), Construction (CON), Wholesale and Retail Trade (WRT), Accommodation and Catering (A&C), Water

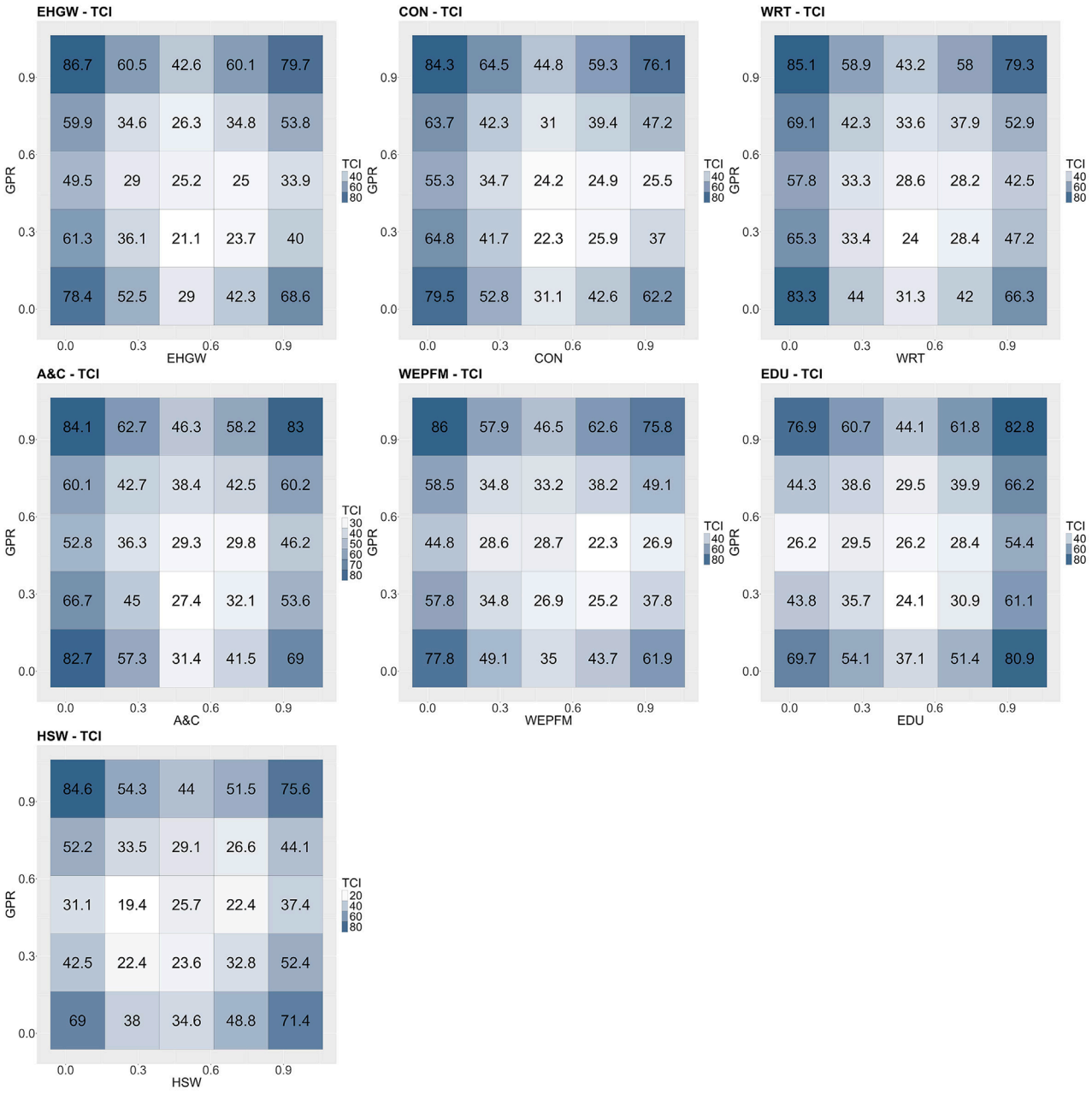


FIGURE 5
Average quantile-on-quantile connectedness between GPR and the tail risk of the modern service industries.

Conservancy, Environment, and Public Facilities Management (WEPFM), Education (EDU), and Health and Social Work (HSW). This structured classification not only clarifies how different industries interact within the economy but also strengthens the analysis of how geopolitical shocks propagate through these interconnected sectors.

We use a QQ connectedness model with a 24-month rolling window, a lag length of one based on the Bayesian Information Criterion, and 20-step-ahead forecasts. In each heat map, the left and right vertical axes represent quantile distribution levels, while the color bar shows total or net spillover intensity. Warmer shades signal stronger connectedness among variables. By examining diverse quantiles, we detect asymmetric spillover effects. Concurrently, time-series plots position time on the horizontal axis, unveiling the temporal evolution of spillovers and distinguishing behaviors in both positive and negative tail risks.

4.1 Quantile-on-quantile total connectedness analysis (TCI)

The TCI measures the connectedness between GPR and the tail risk of 18 industries in the Chinese stock market

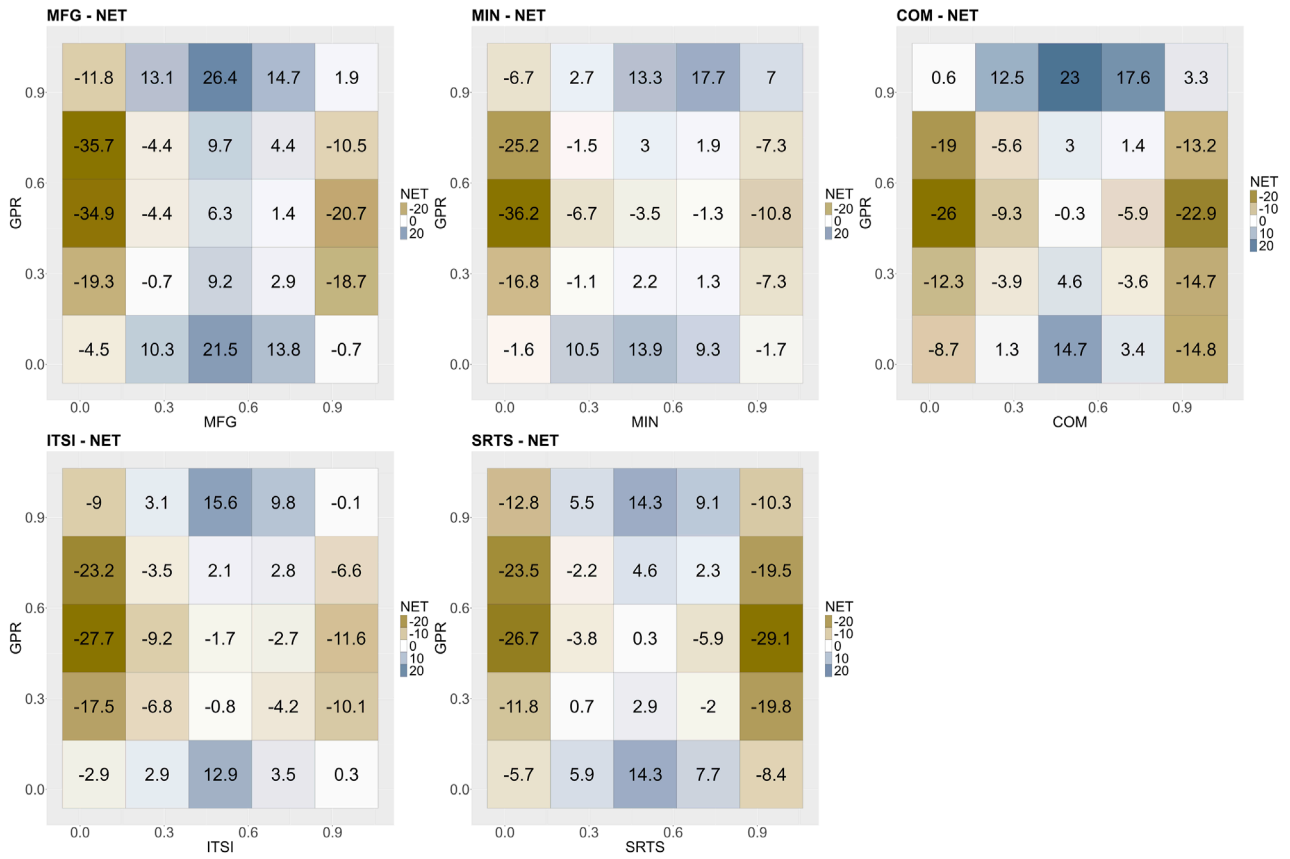


FIGURE 6
Net quantile-on-quantile connectedness between GPR and the tail risk of the dominant industries.

under varying quantile conditions. By examining TCI at multiple quantile levels, we gain deeper insights into how geopolitical risk shapes extreme market outcomes under both mild and acute stress scenarios. Specifically, we focus on the extreme quantiles (τ = 5% and τ = 95%) to capture the upper and lower bounds where geopolitical shocks may induce pronounced tail interdependencies.

Within the dominant industries group (MFG, MIN, COM, ITSI, and SRTS), as depicted in Figure 3, we observe notably elevated TCI values at extreme quantiles. MFG exhibits TCI values of 87.9 and 78.9 at τ = 5% and τ = 95%, respectively, surpassing other industries due to its profound integration with global supply chains. Once geopolitical factors exert negative impacts on international trade via tariffs or export embargoes, manufacturing typically experiences amplified tail volatility. MIN follows a comparable pattern, with TCI values of 83 and 77.9 under extreme conditions, aligning with the notion that mineral and energy resources are strategically sensitive to global political turmoil. Meanwhile, COM sometimes surpasses MFG at specific extreme quantiles, such as 90.2 and 83.3, reflecting the diversified business structure that allows risks to accumulate across multiple segments. ITSI and SRTS exhibit TCI values ranging from 82.8 to 85.3, reflecting their vulnerability to intensified global IT competition, external blockades, or policy frictions in critical research areas, thereby heightening their tail risk exposure. These findings suggest that highly globalized or resource-dependent industries may benefit from tailored stress-testing mechanisms to counter trade barriers and sanctions [18]. Moreover, preempting conflicts by enhancing supply chain diversification or reducing leverage risk emerges as a prudent strategy [20].

For the basic support industries group (AFAHF, CSE, FIN, TSPS, LBS, and RE), as shown in Figure 4, TCI values remain moderately elevated, typically spanning 75–85 at extreme quantiles. Notably, AFAHF can reach 85–85.4 at τ = 5% and τ = 95%, underscoring pronounced tail linkages when agricultural products and trade policies serve as geopolitical bargaining chips. Similarly, CSE reaches 86.9 and 81.7 in tail scenarios, indicating that policy frictions or ideological clashes can abruptly impair cultural exports, intellectual property transactions, and cross-border activities. FIN sustains TCI values of 80.6 and 77.5, a relatively tempered range, yet demonstrates capacity for rapid risk repricing via global capital flows. TSPS occasionally ascends to 86.6 and 77.4, mirroring vulnerabilities in international shipping routes and sentiment-driven freight rate volatility. LBS and RE exhibit analogous fluctuations, hovering between 83 and 88, stemming from their dependence on foreign investment, cross-border mergers, and financing channels, which can shift dramatically amid escalating geopolitical tensions. These observations imply that finance, logistics, and real estate sectors may propagate geopolitical turbulence into broader market volatility. Regulators might thus prioritize monitoring cross-border capital

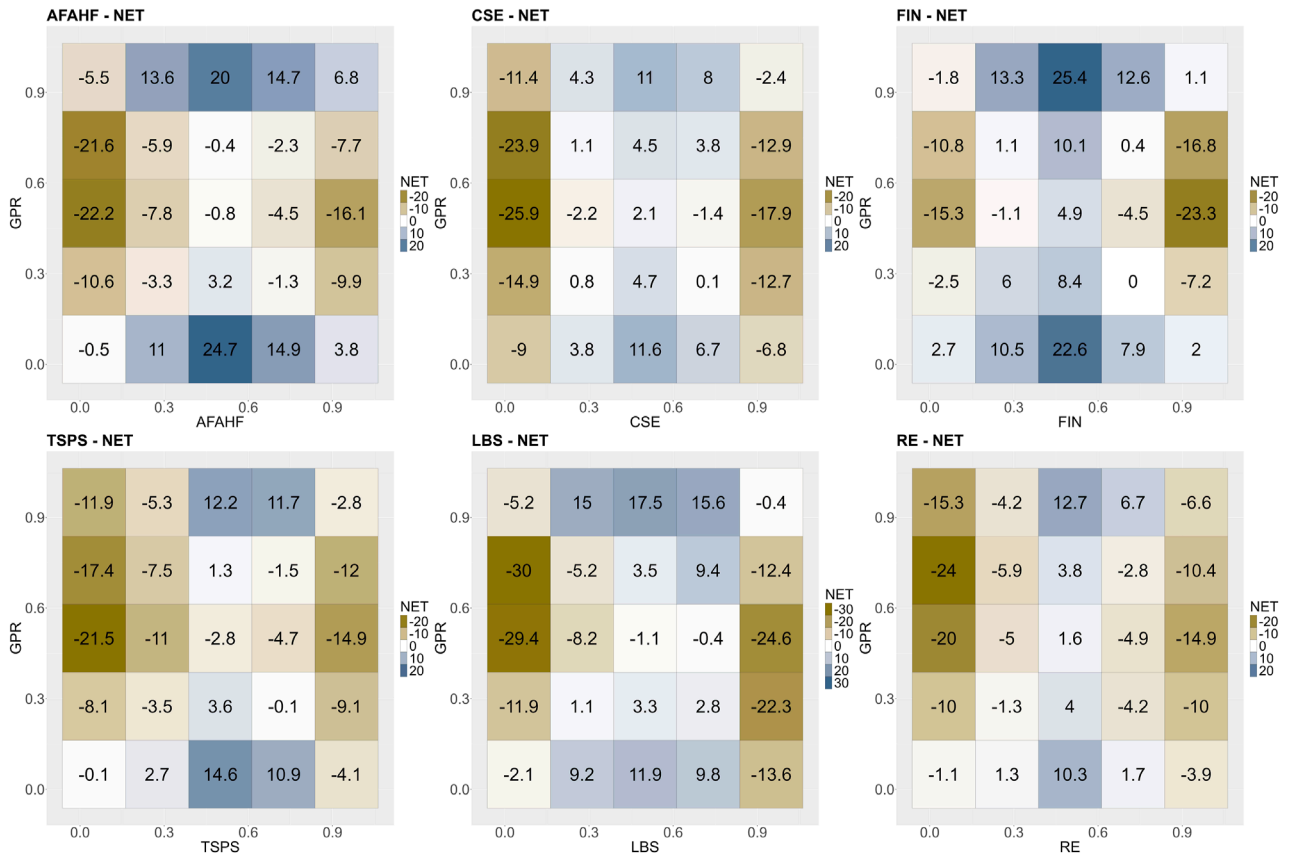


FIGURE 7
Net quantile-on-quantile connectedness between GPR and the tail risk of the basic support industries.

flows and high foreign-investment projects, implementing dynamic margin or liquidity measures during rapid GPR escalations to curb shock amplification [19].

Lastly, the modern service industries group (EHGW, CON, WRT, A&C, WEPFM, EDU, and HSW) generally manifests lower TCI levels relative to other groups. As illustrated in Figure 5, EHGW attains 86.7 and 79.7 at $\tau = 5\%$ and $\tau = 95\%$, evidencing downstream ripple effects from international energy constraints. CON records TCI values of 84.3 and 76.1 under extreme conditions, partly influenced by domestic policy cycles and geopolitical uncertainties in overseas infrastructure ventures. WRT achieves 85.1 and 79.3 at extreme quantiles, propelled by abrupt supply chain disruptions and consumer sentiment shifts during exceptional events. Conversely, A&C, WEPFM, EDU, and HSW typically register TCI values below 75 at median quantiles, suggesting reduced sensitivity to cross-border shocks. Although less reactive to external perturbations, these industries can still amplify market volatility under extreme conditions via demand and cost channels. To mitigate potential tail risk transmission, regulators could bolster resilience through tax incentives, liquidity support, or digital transformation initiatives, while ensuring coordination with traditional sectors to prevent localized constraints from escalating into systemic spillovers [23].

In summary, we identify a distinctive "U-shaped" quantile dependence, with the left ($\tau = 5\%$) and right ($\tau = 95\%$) tails exhibiting peak sensitivity to GPR, while median quantiles remain comparatively stable. Additionally, a slight asymmetry arises when some industries exhibit a more pronounced jump in the left tail than in the right tail. Consequently, relying solely on average or median metrics risks underestimating extreme risk vulnerability [21]. Geopolitical risk exerts a significant amplifying effect under extreme market conditions, with the degree of impact contingent on each industry's level of internationalization, resource dependency, and supply chain geopolitical exposure.

4.2 Net quantile-on-quantile total directional connectedness analysis (NET)

To elucidate the direction of risk flow between GPR and the tail risks of 18 industries, we employ net quantile-on-quantile total directional connectedness (NET), where NET>0 denotes a net "risk sender" and NET<0 denotes a net "risk receiver." Empirically, the largest NET values often appear when one variable is at an extreme quantile (5% or 95%) and the other is at a median quantile (around 50%), or *vice versa*.

In the dominant industries group, as illustrated in Figure 6, MFG can achieve a net value as low as −35.7, whereas MIN can drop to −36.2. When global political tensions escalate, MFG and MIN face immediate exposure to sanctions, raw material supply disruptions, and energy disputes, rendering them

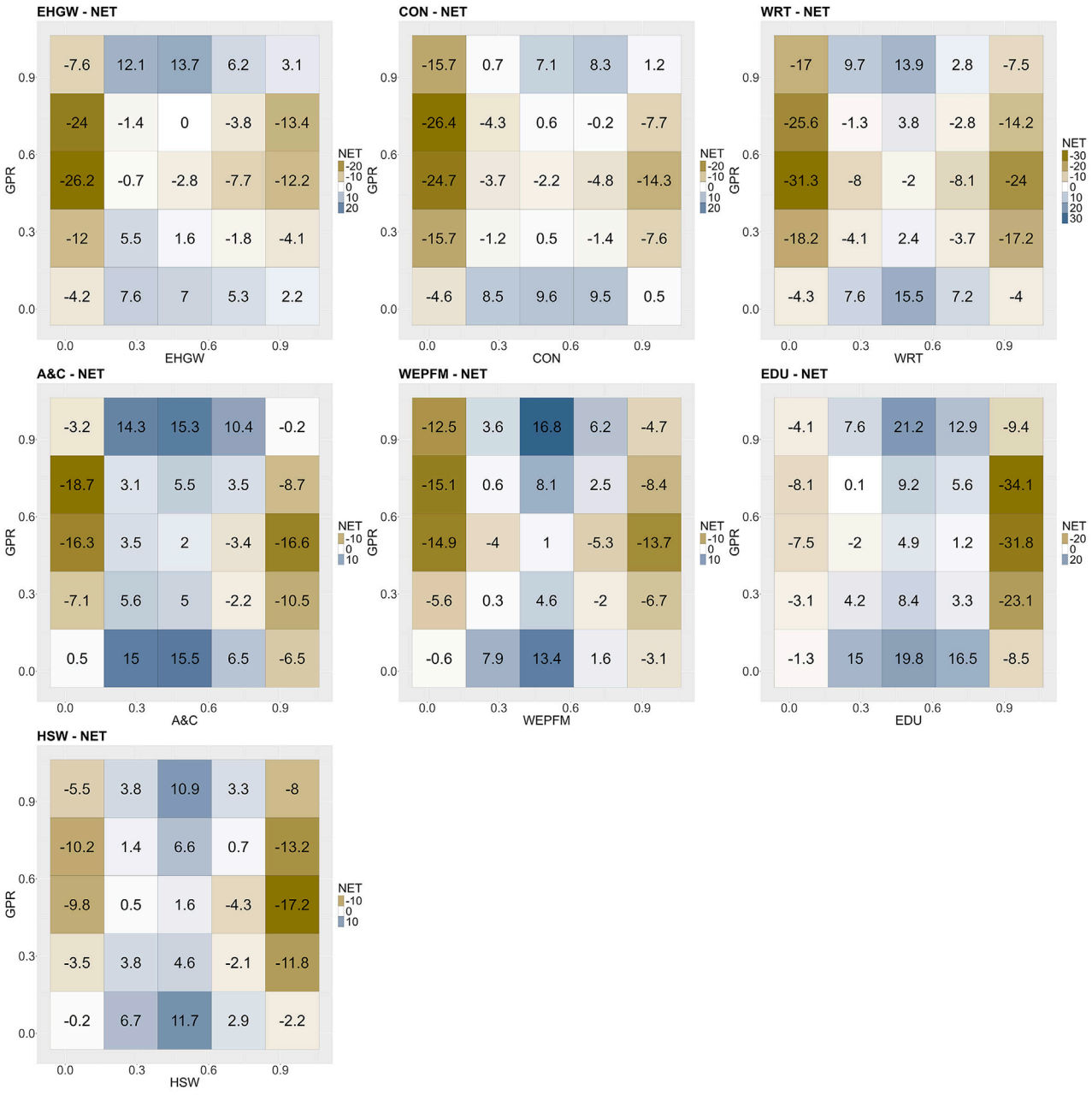


FIGURE 8
Net quantile-on-quantile connectedness between GPR and the tail risk of the modern service industries.

predominantly net risk receivers due to challenges in reallocating or transferring such shocks. Most industries in this group exhibit markedly negative NET values, indicating a strong tendency to absorb risks. For instance, COM consistently displays values around −26, reflecting the inherent challenges of mitigating geopolitical shocks across its multifaceted operations. ITSI and SRTS record NET values around −27.7 and −29.1, respectively, illustrating asymmetric global competition in intellectual property, technology exports, and advanced equipment, which constrains their ability to offload these shocks. Industries with persistent negative NET values are vulnerable to geopolitical and supply chain shocks, warranting proactive regulatory measures—such as intensified foreign investment scrutiny or strategic resource reserves in resource-reliant sectors [22]. In IT and R&D industries, governmental support for research and innovation could reduce external dependency and enhance resilience.

In the basic support industries group, as shown in Figure 7, NET values generally range between −15 and −25, although positive values do appear in certain quantiles. AFAHF, for example, reaches −22.2 at midrange quantiles, showing its vulnerability when agricultural products become bargaining chips in trade disputes. CSE exhibits a pronounced negative value of −25.9, driven in part by

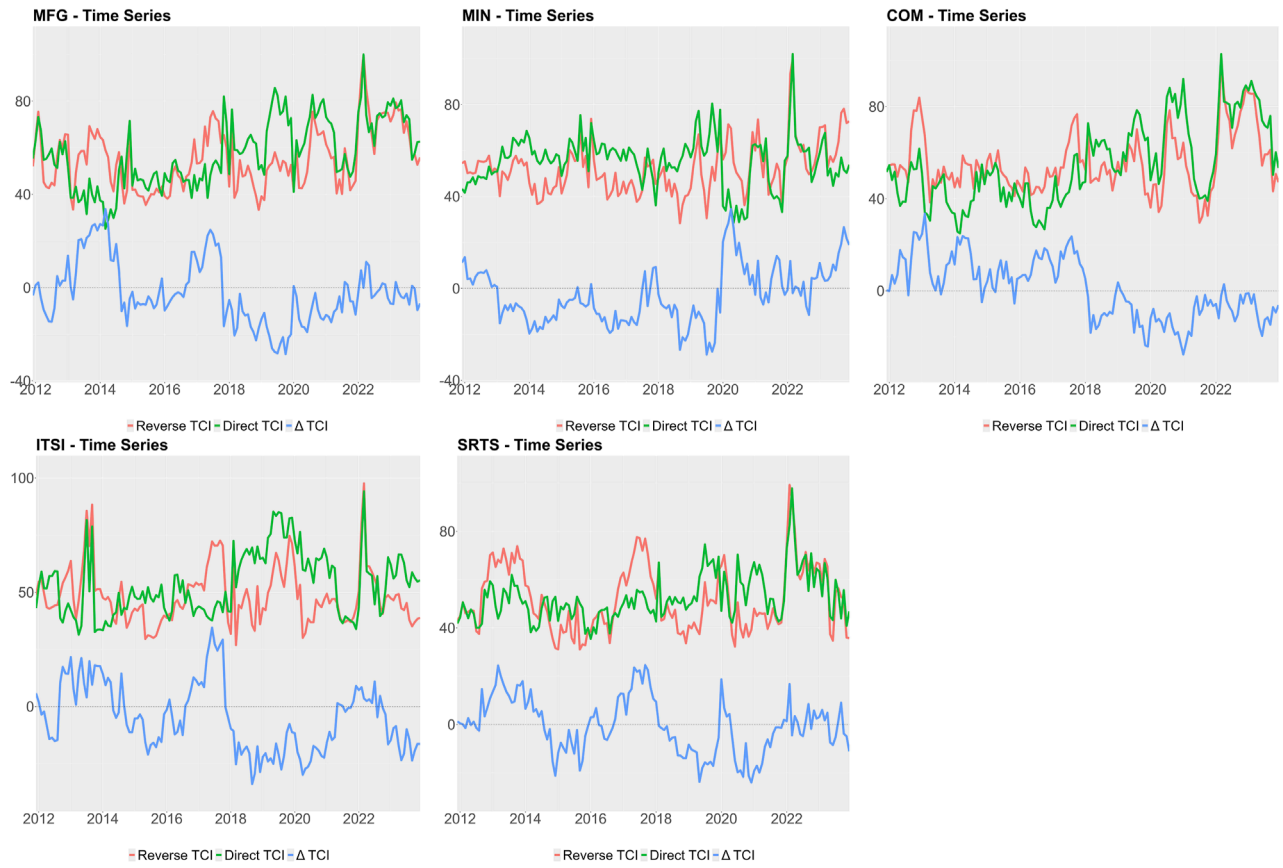


FIGURE 9
Dynamic quantile-on-quantile connectedness between GPR and the tail risk of the dominant industries.

stalled cross-border cultural exchanges and entertainment exports during periods of geopolitical tension. Conversely, FIN occasionally yields positive NET values, such as +22.6, reflecting the leverage and risk transfer mechanisms embedded within financial markets—e.g., heightened speculation or credit contraction can rapidly channel external risks to the broader market. TSPS, LBS, and RE mostly range from −15 to −24, as they passively absorb logistic bottlenecks or capital flow constraints; however, policy support and partial buffers can mitigate some of the tail effects in RE. These findings imply that sectors like finance may not only receive but also send risks under certain stress conditions, emphasizing the need for vigilant cross-industry regulation of leverage cycles [42]. At the same time, strengthening transparency in financial reporting helps prevent systemic spillovers from escalating [43].

Within the modern service industries group, as indicated in Figure 8, NET values are generally smaller in absolute magnitude, around −15 or above, indicating comparatively weaker ties to immediate geopolitical shocks and suggesting that these sectors primarily encounter risk through second-round demand or policy effects. EHGW can reach −26.2 in certain extreme scenarios, but typically remains between −10 and −20, reflecting its passive absorption of higher input costs when global energy prices surge. CON and WRT similarly fluctuate between −15 and −25. Interestingly, A&C sporadically exhibits positive NET values, e.g., 15.5, indicating that sudden contractions in consumer demand can propagate upward along supply chains, effectively making A&C a net sender under specific conditions. Meanwhile, WEPFM, EDU, and HSW remain within ±15, underscoring the domestically oriented, public-service nature of these sectors. This implies that certain modern service industries not only absorb risks but may also transmit shocks upstream. Regulators should focus on these sectors' financing structures and exposures, using flexible contingency plans to mitigate demand-side shocks [21, 23].

NET analysis reveals that most industries—especially those with extensive global exposure, technological barriers, or resource demands—tend to act as net receivers under geopolitical stress. However, sectors like FIN and select consumer industries may become net senders in specific contexts, amplifying market volatility. This pattern intensifies when GPR hits extreme highs or lows while industry risks remain moderate, introducing abrupt external shocks that demand heightened firm- and investor-level attention. Policymakers might thus consider real-time stress tests for globally exposed sectors and targeted capital constraints in financial intermediaries to curb net risk transmission and fortify stability [21, 23].

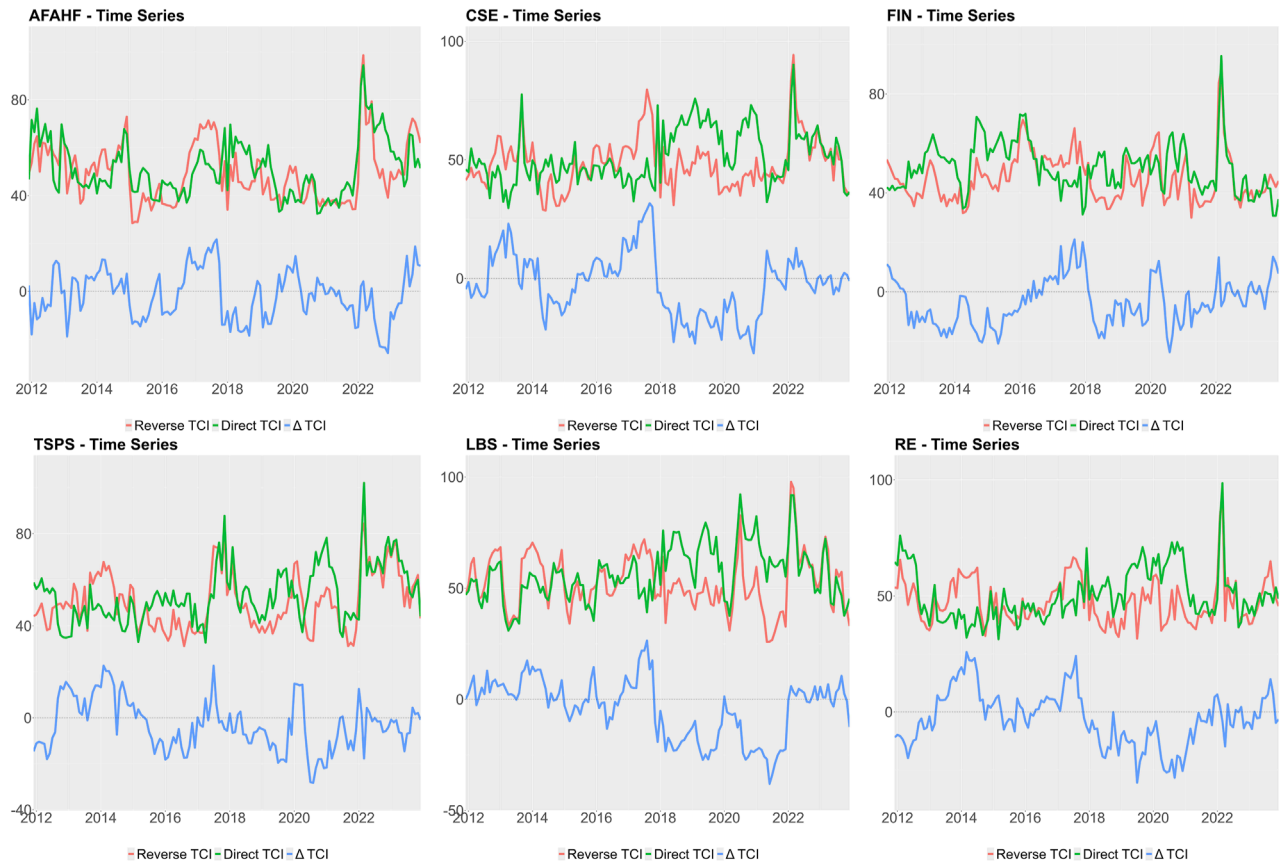


FIGURE 10
Dynamic quantile-on-quantile connectedness between GPR and the tail risk of the basic support industries.

4.3 Dynamic correlation total connectedness analysis (DYNAMIC)

Dynamic Correlation Total Connectedness (DYNAMIC) analysis traces the temporal interplay between GPR and industrial tail risks, spotlighting how major geopolitical events reshape risk network structures and offering deeper insights into connectedness drivers. We present three dynamic connectedness metrics—reverse TCI, direct TCI, and their adjusted difference—alongside their time-varying trajectories.

The dominant industry group exhibits the most striking dynamic connectedness fluctuations, peaking during pivotal geopolitical events. As shown in Figure 9, since the 2018 China–U.S. trade war's onset, MFG has sustained elevated connectedness, surging further during the 2020 COVID-19 outbreak and the 2022 Russia–Ukraine conflict. This aligns with pronounced peaks from 2018 to 2020, followed by a sharp rise circa 2022, fueled by intensified global supply chain strains and commodity price volatility. MIN experiences similarly pronounced fluctuations from 2020 to 2022, coinciding with the Russia–Ukraine-triggered global energy crisis and oil price shocks. COM reacts strongly to key events, particularly during the 2020 pandemic and 2022 conflict escalation, reflecting large diversified firms' exposure to systemic disruptions. Following the 2018 trade disputes and technical embargoes, ITSI has maintained persistently elevated levels of volatility, peaking again in 2020 and 2022, as technology increasingly takes center stage in geopolitical contention. Similarly, SRTS mirrors ITSI's trend, with 2022s global tension spike amplifying tail risk connectedness, driven by uncertainties in cross-border R&D and technology export controls. These patterns mirror broader global transitions, where manufacturing and technology have become focal points in geopolitical competitions [18]. Consequently, Chinese policymakers might benefit from reinforcing cooperative frameworks or setting up cross-border risk-sharing arrangements to mitigate the impact of supply chain disruptions.

Basic support industries exhibit moderate dynamic connectedness, yet display significant peaks around major geopolitical or macroeconomic disruptions. As shown in Figure 10, AFAHF escalates sharply since 2018, peaking during the 2018–2019 China–U.S. agricultural disputes and 2022s Russia–Ukraine-induced grain price surges. CSE spikes in 2018–2019 and 2022, often tied to cross-border copyright tensions or cultural export restrictions. FIN tracks global financial volatility, cresting during the 2015 Chinese stock market turmoil, 2018 trade clashes, 2020 pandemic, and 2022 Russia–Ukraine conflict. TSPS shows heightened connectedness in 2020 lockdowns and 2022 shipping disruptions, while LBS and RE surge from 2020 to 2022, partly due to external financial conditions and domestic policy shifts (e.g., real estate regulations). These peaks suggest that basic support industries' volatility could amplify systemic contagion under external shocks,



FIGURE 11
Dynamic quantile-on-quantile connectedness between GPR and the tail risk of the modern service industries.

especially when global supply-demand and financing channels tighten. Regulators might reinforce resilience via liquidity support in agriculture, logistics, and utilities, alongside tighter oversight of overseas-financed sectors to limit spillovers [19].

Though the modern service industries remain relatively stable compared to other groups, they exhibit sharp, transient peaks. As depicted in Figure 11, EHGW rises notably from 2021 to 2022, lagging upstream mining but reflecting cost shocks from international energy prices. CON, sensitive to domestic real estate cycles, peaks locally in 2022 amid geopolitical shifts and the strict pandemic measures. WRT, A&C, WEPFM, EDU, and HSW spike intensely yet briefly in early 2020 due to COVID-19, reverting to lower volatility by 2021, highlighting reliance on domestic demand and policy support over international factors. Despite milder fluctuations, these sectors face disruptive shocks during global or domestic crises, as seen in early 2020s acute peaks. Temporary support measures—such as credit lines, tax breaks, and strategic digital investments—can help offset short-term cost pressures and

stabilize demand [23]. Coordination with other sectors is critical to prevent localized disruptions from broadening into systemic risks.

In summary, we have three key findings: First, a "U-shaped" quantile dependence prevails between GPR and Chinese stock market tail risks, with TCI surging at $\tau = 5\%$ and $\tau = 95\%$ while mid-quantiles stabilize, underscoring linear correlation's inadequacy for tail dynamics. Second, geopolitical shocks manifest asymmetrically, concentrating impacts during high-risk or stressed periods: globally integrated, resource-intensive sectors (MFG, MIN, ITSI) predominantly receive risks, while FIN occasionally sends risks, heightening market turbulence. Third, dynamic connectedness peaks during seminal events—the 2018 China–U.S. trade war, 2020 COVID-19 crisis, and 2022 Russia–Ukraine conflict—marking turning points in the GPR–industry risk network, with rising, frequent spikes since 2018 signaling growing vulnerability to geopolitical uncertainty.

4.4 Robustness tests

In this study, we performed multiple robustness tests to validate our findings. First, we applied the spillover measurement approach within the traditional generalized vector autoregression (VAR) framework of Diebold and Yilmaz [38], affirming robust connectivity between GPR and industrial tail risks (Supplementary Table A1). Replacing the SJC-copula function with the Clayton copula function [44] preserved stable connectedness (Supplementary Figures A1–A9). When we substituted the US-centric geopolitical risk index with the Chinese geopolitical risk index (GPRC_CHN), the results also remained consistent (Supplementary Figures A10–A18). These methodological variations reinforce the reliability and consistency of our conclusions across diverse paradigms.

Second, we adjusted forecasting horizons (n_fore) and rolling window sizes (window.size), as detailed in the appendix (Supplementary Figures A19–A63). Horizons shifted from 20 months to 10 and 30 months, and window lengths varied from 24 months to 12, 18, and 36 months, consistently upholding the quantile-dependent connectedness of GPR and tail risk across different time horizons. Meanwhile, quantile partitions were also altered from (0.05, 0.25, 0.50, 0.75, 0.95) to (0.10, 0.30, 0.50, 0.70, 0.90), with heightened connectivity at extreme quantiles remaining evident (Supplementary Figures A64–A72). This sustained non-linear tail dependence further validates our primary findings under alternative quantile configurations.

Finally, we analyzed the post-March 2020 period to gauge the pandemic's influence on QQ connectedness (Supplementary Figures A73–A81). Early pandemic phases revealed intensified extreme risk fluctuations, reflecting market actors' heightened sensitivity to geopolitical uncertainty amid global instability. As recovery progressed, tail sensitivity persisted, particularly at higher risk quantiles, confirming rapid risk escalation under uncertainty. Collectively, these tests—spanning alternative methodologies (e.g., copula functions, GVAR), GPR proxies (e.g., GPRC_CHN), and varying horizons, windows, and quantiles—affirm the robustness of our core conclusions.

5 Conclusions, policy implications, limitations and future directions

5.1 Conclusion

The study employs the QQ connectedness approach to investigate the intricate relationship between GPR and tail risk across 18 industries of the Chinese stock market, yielding three key findings. First, the mutual linkage between GPR and tail risk in China demonstrates a pronounced U-shaped quantile dependence, with connectedness intensifying sharply at the lower and upper extremes ($\tau = 5\%$ and $\tau = 95\%$) yet remaining relatively moderate around the median. This underscores a strong nonlinear pattern that traditional linear correlation models fail to capture. Second, the net connectedness analysis reveals significant asymmetry in risk transmission, particularly under extreme conditions marked by heightened geopolitical tensions or market stress. In such scenarios, industries like MFG, MIN, ITSI, and SRTS typically act as "risk receivers," while FIN may, under specific circumstances, emerge as a "risk sender." Finally, dynamic analysis identifies major geopolitical events—such as the 2018 China-U.S. trade war, the 2020 COVID-19 pandemic, and the 2022 Russia-Ukraine conflict—as pivotal turning points in the risk network.

5.2 Policy implications

Based on these empirical findings, the study proposes several targeted policy implications. First, regulatory bodies should establish a quantile-specific risk monitoring framework, paying particular attention to market reactions at the 5% and 95% extremes of geopolitical risk. Specifically, for dominant sectors—MFG, MIN, ITSI, and SRTS—authorities should implement preventive safeguards, for example, timely liquidity injections, before trade tensions escalate. For basic support industries such as FIN, TSPS, and RE, it is vital to monitor and mitigate potential spillover effects in order to avoid amplifying risks during geopolitical conflicts. Meanwhile, modern service industries warrant comparatively moderate regulation to encourage their stabilizing role under conditions of rising domestic demand. Second, investors should leverage the NET analysis, which suggests steering clear of sectors with heightened net risk reception, such as MFG and MIN, during periods of intensifying geopolitical risk, while remaining vigilant about possible risk spillovers from FIN under specific scenarios. Portfolio construction should fully account for the time-varying dynamics highlighted by the DYNAMIC analysis, especially the generally strengthening connectedness between GPR and various industries. Finally, government agencies should strengthen international coordination and communication channels to reduce information asymmetry and mitigate the detrimental impact of geopolitical conflicts on market sentiment. Concurrently, deepening multilateral initiatives—such as the Belt and Road—can broaden access to global markets, thereby bolstering overall economic resilience and enhancing the capacity to navigate geopolitical risks effectively.

5.3 Limitations and future research directions

Despite the robust findings, this study has several limitations. First, our analysis relies on monthly data, which may not fully capture short-term market reactions to sudden geopolitical events. Second, the SJC-copula approach, while effective for tail dependence, may not account for all forms of nonlinear interactions. Third, our focus on China limits the generalizability to other emerging markets with different institutional structures. Future research could address these limitations by: (1) incorporating high-frequency data to capture immediate market responses; (2) exploring alternative nonlinear methodologies such as wavelet analysis or machine learning approaches to enhance model flexibility; (3) extending the framework to compare multiple emerging economies for cross-country validation; (4) integrating investor sentiment indicators to disentangle behavioral factors from fundamental reactions; and (5) developing forward-looking risk measures that combine GPR with alternative data sources such as social media sentiment or satellite imagery to improve predictive capabilities in tail risk management.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Data available only on authorization. Requests to access these datasets should be directed to An Li, lianacd@126.com.

Author contributions

JL: Visualization, Software, Data curation, Conceptualization, Writing – original draft, Methodology, Formal Analysis, Validation, Writing – review and editing. YL: Data curation, Writing – review and editing, Resources, Conceptualization, Methodology, Validation, Formal Analysis, Supervision. FL: Project administration, Funding acquisition, Resources, Validation, Methodology, Supervision, Conceptualization, Software, Writing – review and editing. AL: Data curation, Visualization, Software, Supervision, Project administration, Formal Analysis, Conceptualization, Writing – review and editing, Methodology.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study is supported by the National Natural Science Foundation of China (Grant no. 72463032).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1612695/full#supplementary-material

References

1. Boubaker S, Goodell JW, Pandey DK, Kumari V. Heterogeneous impacts of wars on global equity markets: evidence from the invasion of Ukraine. *Finance Res Lett* (2022) 48:102934. doi:10.1016/j.frl.2022.102934

2. Caldara D, Iacoviello M. Measuring geopolitical risk. *Am Econ Rev* (2022) 112:1194–225. doi:10.1257/aer.20191823

3. Zhang Y, Zhang Y, Ren X, Jin M. Geopolitical risk exposure and stock returns: evidence from China. *Finance Res Lett* (2024) 64:105479. doi:10.1016/j.frl.2024.105479

4. Pan C, Zhang W, Wang W. Global geopolitical risk and volatility connectedness among China's sectoral stock markets. *Finance Res Lett* (2023) 58:104487. doi:10.1016/j.frl.2023.104487

5. Adrian T, Brunnermeier MK. CoVaR. *Am Econ Rev* (2016) 106:1705–41. doi:10.1257/aer.20120555

6. Giglio S, Kelly B, Pruitt S. Systemic risk and the macroeconomy: an empirical evaluation. *J Financial Econ* (2016) 119:457–71. doi:10.1016/j.jfineco.2016.01.010

7. Bai J, Perron P. Computation and analysis of multiple structural change models. *J Appl Econom* (2003) 18:1–22. doi:10.1002/jae.659

8. Baur DG, Lucey BM. Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Rev* (2010) 45:217–29. doi:10.1111/j.1540-6288.2010.00244.x

9. Guo P, Shi J. Geopolitical risks, investor sentiment and industry stock market volatility in China: evidence from a quantile regression approach. *North Am J Econ Finance* (2024) 72:102139. doi:10.1016/j.najef.2024.102139

10. Baker SR, Bloom N, Davis SJ. Measuring economic policy uncertainty. *Q J Econ* (2016) 131:1593–636. doi:10.1093/qje/qjw024

11. Ahlgren N, Antell J. Stock market linkages and financial contagion: a cobreaking analysis. *Q Rev Econ Finance* (2010) 50:157–66. doi:10.1016/j.qref.2009.12.004

12. Mollah S, Quoreshi AMMS, Zafirov G. Equity market contagion during global financial and eurozone crises: evidence from a dynamic correlation analysis. *J Int Financial Markets Institutions Money* (2016) 41:151–67. doi:10.1016/j.intfin.2015.12.010

13. Gabauer D, Stenfors A. Quantile-on-quantile connectedness measures: evidence from the US treasury yield curve. *Finance Res Lett* (2024) 60:104852. doi:10.1016/j.frl.2023.104852

14. Naifar N. Monetary policy expectations and financial markets: a quantile-on-quantile connectedness approach. *North Am J Econ Finance* (2025) 77:102389. doi:10.1016/j.najef.2025.102389

15. Niu H, Zhang S. Asymmetric effects of commodity and stock market on Chinese green market: evidence from wavelet-based quantile-on-quantile approach. *Renew Energ* (2024) 230:120794. doi:10.1016/j.renene.2024.120794

16. Naeem MA, Pham L, Senthilkumar A, Karim S. Oil shocks and BRIC markets: evidence from extreme quantile approach. *Energ Econ* (2022) 108:105932. doi:10.1016/j.eneco.2022.105932

17. NguyenHuu T, Örsal DK. Geopolitical risks and financial stress in emerging economies. *World Economy* (2024) 47:217–37. doi:10.1111/twec.13529

18. Umar Z, Hadad E, Phiri A, Teplova T. Dynamics of asymmetric connectedness among magnificent seven technology giants: insights from QVAR analysis. *Q Rev Econ Finance* (2025) 101:101977. doi:10.1016/j.qref.2025.101977

19. Bulut E, Marangoz C. Exploring the impact of economic recession indicators on global financial markets: a QVAR analysis. *Int Rev Financial Anal* (2025) 99:103966. doi:10.1016/j.irfa.2025.103966

20. Hoque ME, Billah M, Kapar B, Naeem MA. Quantifying the volatility spillover dynamics between financial stress and US financial sectors: evidence from QVAR connectedness. *Int Rev Financial Anal* (2024) 95:103434. doi:10.1016/j.irfa.2024.103434

21. Patra S, Malik K. Return and volatility connectedness among US and Latin American markets: a QVAR approach with implications for hedging and portfolio diversification. *Glob Finance J* (2025) 65:101094. doi:10.1016/j.gfj.2025.101094

22. Ren X, Wang S, Mao W, Gozgor G. Greening the energy industry: an efficiency analysis of China's listed new energy companies and its market spillovers. *Energ Econ* (2025) 145:108414. doi:10.1016/j.eneco.2025.108414

23. Shu M, Liu B, Ouyang W, Sun R, Lin Y. Multi-scale dynamic correlation and information spillover effects between climate risks and digital cryptocurrencies: based on wavelet analysis and time-frequency domain QVAR. *Physica A: Stat Mech Appl* (2025) 663:130443. doi:10.1016/j.physa.2025.130443

24. Pástor Ľ, Veronesi P. Political uncertainty and risk premia. *J Financial Econ* (2013) 110:520–45. doi:10.1016/j.jfineco.2013.08.007

25. Hoque ME, Zaidi MAS. Global and country-specific geopolitical risk uncertainty and stock return of fragile emerging economies. *Borsa Istanbul Rev* (2020) 20:197–213. doi:10.1016/j.bir.2020.05.001

26. Triki MB, Ben Maatoug A. The GOLD market as a safe haven against the stock market uncertainty: evidence from geopolitical risk. *Resour Policy* (2021) 70:101872. doi:10.1016/j.resourpol.2020.101872

27. Huang Z, Ye C, Lai F, Li Y, Li X. Dancing between threats and conflicts: how Chinese energy companies invest amidst global geopolitical risks. *Energ Econ* (2024) 139:107907. doi:10.1016/j.eneco.2024.107907

28. Lai F, Xiong D, Zhu S, Li Y, Tan Y. Will geopolitical risks only inhibit corporate investment? Evidence from China. *Pacific-Basin Finance J* (2023) 82:102134. doi:10.1016/j.pacfin.2023.102134

29. Lai F, Li S, Lv L, Zhu S. Do global geopolitical risks affect connectedness of global stock market contagion network? Evidence from quantile-on-quantile regression. *Front Phys* (2023) 11:1124092. doi:10.3389/fphy.2023.1124092

30. Gong X-L, Feng Y-K, Liu J-M, Xiong X. Study on international energy market and geopolitical risk contagion based on complex network. *Resour Policy* (2023) 82:103495. doi:10.1016/j.resourpol.2023.103495

31. Zhao Y, Chen L, Zhang Y. Spillover effects of geopolitical risks on global energy markets: evidence from CoVaR and CAViaR-EGARCH model. *Energy Exploration Exploitation* (2024) 42:772–88. doi:10.1177/01445987231196617

32. Gulen H, Ion M. Policy uncertainty and corporate investment. *Rev Financ Stud* (2015) 2015:hhv050. doi:10.1093/rfs/hhv050

33. Kelly B, Pástor Ľ, Veronesi P. The price of political uncertainty: theory and evidence from the option market. *J Finance* (2016) 71:2417–80. doi:10.1111/jofi.12406

34. Balcilar M, Bonato M, Demirer R, Gupta R. Geopolitical risks and stock market dynamics of the BRICS. *Econ Syst* (2018) 42:295–306. doi:10.1016/j.ecosys.2017.05.008

35. Demirer R, Gupta R, Suleman T, Wohar ME. Time-varying rare disaster risks, oil returns and volatility. *Energ Econ* (2018) 75:239–48. doi:10.1016/j.eneco.2018.08.021

36. Li S, Tu D, Zeng Y, Gong C, Yuan D. Does geopolitical risk matter in crude oil and stock markets? Evidence from disaggregated data. *Energ Econ* (2022) 113:106191. doi:10.1016/j.eneco.2022.106191

37. Diebold FX, Yilmaz K. On the network topology of variance decompositions: measuring the connectedness of financial firms. *J Econom* (2014) 182:119–34. doi:10.1016/j.jeconom.2014.04.012

38. Diebold FX, Yilmaz K. Better to give than to receive: predictive directional measurement of volatility spillovers. *Int J Forecast* (2012) 28:57–66. doi:10.1016/j.ijforecast.2011.02.006

39. Chatziantoniou I, Gabauer D, Stenfors A. Interest rate swaps and the transmission mechanism of monetary policy: a quantile connectedness approach. *Econ Lett* (2021) 204:109891. doi:10.1016/j.econlet.2021.109891

40. Pesaran HH, Shin Y. Generalized impulse response analysis in linear multivariate models. *Econ Lett* (1998) 58:17–29. doi:10.1016/S0165-1765(97)00214-0

41. Patton AJ. Modelling asymmetric exchange rate dependence. *Int Econ Rev* (2006) 47:527–56. doi:10.1111/j.1468-2354.2006.00387.x

42. Wang H, Sua LS, Huang J, Ortiz J, Alidaee B. Will southeast asia be the next global manufacturing hub? A multiway cointegration, causality, and dynamic connectedness analyses. *Emerging Markets Rev* (2024) 63:101217. doi:10.1016/j.ememar.2024.101217

43. Shi G, Chen Z, Luo W, Wei Z. Cross-border spillover of imported sovereign risk to China: key factors identification based on XGBoost-SHAP explainable machine learning algorithm. *Finance Res Lett* (2024) 70:106307. doi:10.1016/j.frl.2024.106307

44. Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* (1978) 65:141–51. doi:10.1093/biomet/65.1.141

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 05 June 2025
DOI 10.3389/fphy.2025.1616099

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Hailing Li,
Hunan Agricultural University, China
Xuesong Dong,
Hunan Agricultural University, China
Yingli Li,
Zhengzhou University, China

*CORRESPONDENCE
Hao Nong,
231601011@csu.edu.cn

RECEIVED 22 April 2025
ACCEPTED 16 May 2025
PUBLISHED 05 June 2025

CITATION
Zheng R, Nong H, Chen Q, Shao L, Zhang H and Zhang Y (2025) Study on risk contagion between rare earth and new energy-related metals markets from the perspective of co-production and co-consumption. *Front. Phys.* 13:1616099. doi: 10.3389/fphy.2025.1616099

COPYRIGHT
© 2025 Zheng, Nong, Chen, Shao, Zhang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Study on risk contagion between rare earth and new energy-related metals markets from the perspective of co-production and co-consumption

Renrui Zheng[1,2], Hao Nong[3]*, Qishen Chen[4], Liuguo Shao[3,5], Hongtao Zhang[1] and Yanfei Zhang[1,4]

[1]School of Earth Sciences and Resource, China University of Geosciences (Beijing), Beijing, China, [2]Chief Engineer's Office, China Geological Survey, Beijing, China, [3]School of Business, Central South University, Changsha, China, [4]Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing, China, [5]Institute of Metal Resources Strategy, Central South University, Changsha, China

Introduction: This paper investigates risk contagion between rare earth and new energy-related metal markets, with a focus on higher-order moment risk spillovers from a co-production and co-consumption perspective. Understanding these dynamics is crucial due to the growing interdependence of these markets and their susceptibility to policy shifts and extreme events.

Methods: Using higher-order moment risk measures, time-domain spillover indexes, and frequency-domain spillover indexes, we analyze return, volatility, skewness, and kurtosis risk spillovers across metal markets from 2014 to 2024. The study employs maximum pairwise spillover portfolio analysis to identify key risk transmission channels.

Results: The findings reveal that return and volatility risk spillovers increase over time, while skewness risk exhibits short-term high spillovers and kurtosis risk shows long-term persistence, reflecting market sensitivity to asymmetric information and extreme events. Policy changes and major events (e.g., trade disputes, geopolitical tensions) amplify uncertainty and risk contagion. Markets with stable co-production and co-consumption relationships are identified as major risk sources. Frequency-domain analysis highlights the rare earth market's dominant role in long- and short-term spillovers, with significant external spillover effects.

Discussion: The results underscore the rare earth market's critical influence on new energy metal markets, suggesting that investors should closely monitor its dynamics to optimize risk management. The study provides policymakers and market participants with insights into risk transmission mechanisms and the systemic role of rare earths in the global economy.

KEYWORDS

co-production, co-consumption, rare earths, higher-order moment, risk contagion

1 Introduction

With the rapid development of the clean energy industry, the demand for rare earths and new energy-related metals is growing at a high rate [1]. The impact of the rare earths market on the new energy-related metals market is becoming increasingly prominent, particularly due to the unique applications of rare earths and their increasing price volatility [2,3]. Rare earth price volatility impacts the entire industrial chain and supply chain. It directly affects company costs and profits. Moreover, it can have far-reaching consequences for overall industry development. Currently, China has become the largest producer and consumer of rare earths in the world [4]. The United States, the European Union, Japan, Australia and other countries have included rare earths in their key mineral lists, and competition for the development and use of rare earths and the development of the industry chain is intensifying[1]. Therefore, in-depth research on the price risk spillover effect of rare earths and its impact on the new energy industry, as well as exploring the corresponding risk management strategies, is of great theoretical and practical significance [5]. This study aims to critically analyse the impact mechanism of rare earth price fluctuations on the new energy industry chain, explore risk management strategies, and provide valuable suggestions for the sustainable development of the clean energy sector.

The production-consumption relationship of rare earth elements (REEs) and metals associated with new energy sources has become increasingly complex in the clean energy sector, with the phenomena of co-production and co-consumption in particular requiring attention. Co-production refers to the simultaneous occurrence of these two elements in the production process [6], while co-consumption refers to their use together in new energy products [7]. Co-production and co-consumption make the links in the clean energy industry chain interconnected, and an impact on one link is transmitted to other links, leading to increased uncertainty in market volatility [8–11]. That is, the development of the new energy industry relies on the supply of specific raw materials, and the prices of these raw materials may be affected by fluctuations in the international market; at the same time, the consumption demand for rare earths and new energy-related metals is affected by a variety of factors such as policies, regulations and technological development. Therefore, if price fluctuations or risk spillovers occur in a certain link, they will quickly spread to the whole industrial chain, exacerbating the risk spillover effect of price fluctuations. This complex linkage and transmission effect makes the entire clean energy industry chain more vulnerable to external risks and has far-reaching implications for market volatility.

The risk contagion situation between rare earth and new energy-related metals markets is facing further changes due to policies and events. Policies have a direct impact on the development of the rare earth and new energy-related metals industry, and adjustments in import and export policies, resource extraction management, etc. May lead to adjustments in market supply and demand patterns and price volatility [12]. At the same time, some key events, such as international trade frictions and geopolitical tensions, may also affect the market and exacerbate uncertainty [13–15]. In this context, the risk contagion effect between rare earths and new energy-related metals markets is intensifying, and market participants are facing more complex market environments and risk challenges. Therefore, in-depth research on the mechanisms and factors influencing risk contagion between these two markets should take into account the impact of relevant policies and events.

In summary, this paper aims to answer the following questions: what are the characteristics of the different moments of risk spillovers between rare earth elements and the new energy sector, and what are the relevant influencing policies and events? How does the complexity of production-consumption relationships affect the evolution of pairwise risk spillovers between markets? How do risk spillovers differ in the analysis of market linkages in different frequency domains?

The empirical results show that the risk spillover between return and volatility is persistently high between 2016 and 2024. Whereas skewness risk exhibits a short-term high spillover, kurtosis risk is a long-term high spillover, reflecting the complexity of the market in dealing with asymmetric information and extreme events. When the four types of risk rose significantly, they mostly stemmed from government policy support, resource control or international competition. These policies raise market expectations of future returns while triggering an increase in short-term uncertainty, leading to frequent extreme market volatility.

In the pairwise risk premium evolution, after 2017, with the advancement of technology, changes in market demand and the gradual maturity of the relevant industrial chain, specific metals with common consumption and joint production relationships gradually evolved into a major source of risk. In the skewness and kurtosis risk premium analyses, the maximum risk premium value is significantly higher than return and volatility, suggesting the existence of strong asymmetric price volatility risk-transfer effects in the market as well as abnormal volatility connectivity in the face of external shocks.

In the comparative analysis of risky long- and short-term market correlations for each order of moments, it is found that the rare earth market exhibits a more pronounced pattern of long- and short-term correlations. This suggests that rare earths are not only affected by short-term speculative funds and market news, but also driven by long-term supply and demand, industrial policy and other factors. The existence of long- and short-term correlation patterns is of great significance for risk management and investment decisions in the rare earth market.

The contribution of this paper is: 1. To consider risk contagion between markets from the perspective of the production-consumption relationship of co-production and co-consumption.

1 Details can be found in the report below: Rarw Earths Statistics and Information. https://www.usgs.gov/centers/national-minerals-information-center/rare-earths-statistics-and-information. A Federal Strategy to Ensure Secure and Reliable Supplies of Critical Minerals. https://www.commerce.gov/data-and-reports/reports/2019/06/federal-strategy-ensure-secure-and-reliable-supplies-critical-minerals. Study on the EU's list of Critical Raw Materials Final Report. https://op.europa.eu/en/publication-detail/-/publication/c0d5292a-ee54-11ea-991b-01aa75ed71a1/language-en. Critical Minerals Strategy 2022. https://cdn.prod.website-files.com/6164f987875e87a4dbb1404e/635f07987a08e28cf7d1eee9_2022%20Australian%20critical-minerals-strategy_0-compressed.pdf.

By analysing the price volatility relationship from the raw material side, this approach can capture the market dynamics at the same level in the supply chain in a more scientific way than directly linking the price of rare earth elements to the midstream and downstream industries in the new energy industry. 2. The maximum pairwise spillover index analysis carried out in this paper is able to reveal the interactions between different metals in the evolution of market prices. By examining the maximum pairwise spillover index, we can more clearly identify which materials have significant bilateral relationships with each other and which production-consumption relationships are most susceptible to potential impacts. Through an in-depth study of the maximum pairwise spillover index, we can discover more information about the operating laws of the market and thus provide a more reliable reference for future risk management and decision making. 3. By analysing the evolution of the time-frequency spillover network for different moments of the order of magnitude, it is possible to help investors and policymakers understand the changes in the market structure, and to accurately identify and quantify how rare earth elements influence the price fluctuations of new energy-related metals, thus revealing the price fluctuations, which in turn reveals the transmission mechanism and the time-continuity effects of market risks.

In the remaining chapters of the article, the paper will further demonstrate the significance of this paper by providing a literature review in Chapter 2. The methodology used in this paper is presented in Chapter 3, where readers can deepen their understanding of econometric modelling theory. Chapter 4 contains data selection and descriptive statistical analyses. Detailed empirical analyses will be developed in Chapter 5. Chapter 6 provides a summary of the findings of this paper.

2 Literature review

In recent years, the research literature on rare earths has focused on analysing their importance in financial markets. For example, Baldi et al. [16] show in their paper that there is a negative correlation between rare earths price changes and the stock market performance of some clean energy indices during periods of price increases. Reboredo and Ugolini [17] empirically show that the price linkages between rare earths and other stock markets vary under different volatility mechanisms. Under low volatility rare earth stocks are linked to base metal markets, receiving and transmitting considerable price spillovers. In the high volatility scenario, rare earth prices are more strongly linked to price fluctuations in clean energy, oil and general stock markets.

With the deepening of research, the importance of higher-order moments, time-frequency spillover analysis and its related methods for the analysis of rare-earth market linkages has gradually attracted the attention of scholars. Gao and Liu [18] show empirically that the rare-earth metal (REM) market is a receiver of net spillovers from the base metal, clean energy and ESG markets. Gao et al. [5] reveal a significant higher-order moments correlation between the clean energy and rare-earth markets. In particular, wind and solar exhibit significant net spillovers of skewness and kurtosis to rare earths. These higher-order moment spillovers are time-varying and are exacerbated during major geopolitical crises such as the COVID-19 pandemic and the Russo-Ukrainian war. Ding et al. [19] use a frequency-domain spillover index methodology to find that rare-earth metals are the most susceptible to spillovers from both the clean-energy and the base-metal markets, especially in the long run. Kamal and Bouri [20] used the maximum overlap discrete wavelet transform method to decompose the original return series into different time scales and then analysed the pairwise dependence using vine-copula. The results show that the rare earth market mitigates the dependence between world equities and clean energy markets at different time scales, especially during the COVID-19 epidemic. Zheng et al. [21] examine the volatility spillover effects of the rare earth market with the crude oil, renewable energy, and high-technology markets using a wavelet analysis framework, confirming the importance of rare earths in improving the overall risk transfer of the system. These findings provide new perspectives for understanding the role of rare earths in the global economic system and provide valuable information for relevant policymakers and investors.

It is known from the findings of many scholars that there are differences in the roles of various types of rare earth elements [22–25]. The different production-consumption relationships between rare earths and metal raw materials should be considered in market linkage analyses. Metals used for the same product can be considered as co-consumption metals, and market shocks for this product will affect the price markets of both metals simultaneously. Exploring the linkages within the metals market from a co-consumption perspective can help to get a comprehensive picture of the price volatility patterns in the clean energy metals market, and also help to deepen the understanding of the relationship between the clean energy and metals markets [7]. Regarding co-production, rare earth elements are usually symbiotic, and different rare earth deposits have different rare earth element compositions, which can only be separated through smelting to form different rare earth metals or rare earth oxides [26]. Also rare earths will coexist with minerals such as iron and nickel. This concept emphasises the integrated use of mineral resources, which helps to reduce production costs and increase resource efficiency. Co-production not only makes the production chain more economical, but also promotes environmental protection as it reduces mining pressure on the environment [6]. As a result, the complexity of the two types of production and consumption patterns may make the market volatility and price changes more spillover effects, and the overall impacts of the huge market uncertainty on the industrial chain need to be thoroughly studied and effectively managed [8–11]. It can be seen that many studies mainly examine the metal market as a whole, ignoring the heterogeneity between different clean energy metal markets.

In conclusion, there has been a wealth of research on rare earth markets in recent years, but a number of areas remain worthy of attention. Currently, most of the literature focuses on analysing the overall linkages between the rare-earth market as a whole and other financial markets, and there is a lack of research on the volatility and impacts of individual rare-earth element markets, a perspective that tends to ignore the heterogeneity of the different rare-earth elements and their complexity in co-consumption and co-production with other metals. By analysing the relationship between price fluctuations in terms of raw materials, this approach captures market dynamics at the same level in the supply chain in a more scientific way than by linking rare earth

prices directly to midstream and downstream industries in the new energy sector.

Further, there are still areas worth exploring under the "co-production-co-consumption" perspective. Firstly, relevant studies have mainly focused on the traditional metal market, and the dynamic correlation mechanism between rare earth elements and key metals of new energy sources has not been sufficiently explored. Furthermore, most of the existing literature is limited to the analysis of basic risk indicators such as return and volatility, and fails to examine the cross-market contagion law of extreme risk characteristics such as skewness and kurtosis. Finally, research generally lacks systematic analyses of how policy adjustments and market shocks transmit risk through the co-production-consumption channel, and in particular pays insufficient attention to the dynamic evolution of risk contagion. These limitations make it difficult for existing studies to provide comprehensive and effective theoretical support and practical guidance for risk management in the new energy industry chain under the perspective of "co-production and co-consumption".

In view of the above, this paper analyses the time-frequency risk spillover index under the high-order moment characteristics. In this way, it reveals more comprehensively the risk transmission characteristics between rare earth elements and new energy metals in various events under the perspective of "co-production and co-consumption". At the same time, we consider the central position of the Chinese market in the global supply chain. It is also important to study the linkages between its rare earth market and other raw materials for the new energy industry. By analysing the dynamics and trends of the Chinese market, we can reveal the role of rare earth elements in supporting the new energy industry. And, it will provide a reference for other countries and enterprises to promote the optimal allocation of global resources and the innovative development of new energy technologies.

3 Methodology

3.1 Higher-order moment risk measures

In this paper, the GJRSK model proposed by Nakagawa and Uchiyama [27] will be used to estimate the conditional volatility, conditional skewness, and conditional kurtosis characteristics of the raw materials and rare earth elements markets in the new energy sector. The model allows for asymmetric responses to positive and negative shocks and is constructed as follows (i.e. Equations 1–6):

$$r_t = a_1 r_{t-1} + \varepsilon_t \tag{1}$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \beta_3 \varepsilon_{t-1}^2 I_{(\eta_{t-1}<0)} \tag{2}$$

$$s_t = \gamma_0 + \gamma_1 \eta_{t-1}^3 + \gamma_2 s_{t-1} + \gamma_3 \eta_{t-1}^3 I_{(\eta_{t-1}<0)} \tag{3}$$

$$k_t = \delta_0 + \delta_1 \eta_{t-1}^4 + \delta_2 k_{t-1} + \delta_3 \eta_{t-1}^4 I_{(\eta_{t-1}<0)} \tag{4}$$

$$\eta_t = h_t^{-\frac{1}{2}} \varepsilon_t \tag{5}$$

$$\eta_t | I_{t-1} \sim g(0,1,s_t,k_t) \tag{6}$$

Where r_t denotes the return of each market, calculated as $r_t = \ln(P_t) - \ln(P_{t-1})$. P_t is the price of each market at time t, h_t, s_t and k_t denote conditional volatility, conditional skewness and conditional kurtosis, respectively. I_{t-1} denotes the information set at $t-1$. $g(0,1,s_t,k_t)$ is the probability density function with mean 0, variance 1, skewness s_t and kurtosis k_t. GJRSK GJRSK model can be estimated by maximising the log-likelihood function.

Empirically, the GJRSK model not only measures the magnitude of price rises and falls (e.g., how much a stock has risen or fallen), but also captures extreme changes in market sentiment (e.g., "panic selling" or "excessive speculation"). This helps investors recognise when the market is likely to be volatile or unusual, so they can adjust their strategies in advance.

3.2 Time-domain spillover indexes

The risk spillover analysis in this paper and will be based on the time-varying spillover index, which is an important dynamic connectivity measure proposed by Antonakakis et al. [28] in recent years, which no longer needs to set up a rolling window compared to the previous spillover index, and directly constructs a variance decomposition from the TVP-VAR model to effectively portray the risk spillover effect between different financial markets. Here the TVP-VAR(p) model will be written as (i.e. Equations 7, 8):

$$y_t = A_t Z_{t-1} + \varepsilon_t \quad \varepsilon_t | \Omega_{t-1} \sim N(0, \Sigma_t) \tag{7}$$

$$vec(A_t) = vec(A_{t-1}) + \zeta_t \quad \zeta_t | \Omega_{t-1} \sim N(0, \Xi_t) \tag{8}$$

The TVP-VAR can be transformed into its vector moving average (VMA) representation (i.e. Equations 9, 10):

$$y_t = J' \left(M_t^{k-1} z_{t-k-1} + \sum_{j=0}^{k} M_t^j \eta_{t-j} \right) \tag{9}$$

$$M_t = \begin{pmatrix} A_t & 0_{p \times m} \\ I_{m(p-1)} & 0_{m(p-1) \times m} \end{pmatrix}, \eta_t = (\varepsilon_t, 0, \cdots, 0)', J = (I, 0, \cdots, 0)', z_{t-1} = (y_{t-1}, y_{t-2}, \cdots, y_{t-p})' \tag{10}$$

$A_t = (A_{1t}, A_{2t}, \cdots, A_{pt})'$. M_t is an $mp \times mp$ matrix, η_t is an $mp \times 1$ matrix, and J is an $mp \times m$ matrix. As k approaches ∞, Equation 9 can be written as Equations 11, 12:

$$y_t = \lim_{k \to \infty} J' \left(M_t^{k-1} z_{t-k-1} + \sum_{j=0}^{k} M_t^j \eta_{t-j} \right) = \sum_{j=0}^{\infty} J' M_t^j \eta_{t-j} \tag{11}$$

$$y_t = \sum_{j=0}^{\infty} J' M_t^{k-1} J \varepsilon_{t-j}, B_{jt} = J' M_t^j J, j = 0, 1, \cdots, y_t = \sum_{j=0}^{\infty} B_{jt} \varepsilon_{t-j} \tag{12}$$

B_{jt} is an $m \times m$ matrix. $GIRFs(\Psi_{ij,t}(H))$ represent the responses of all variables j following a shock in variable i. The differences between an H -step-ahead forecast can be calculated by Equations 13–15:

$$GIRF_t(H, \delta_{j,t}, \Omega_{t-1}) = E(y_{t+H} | e_j = \delta_{j,t}, \Omega_{t-1}) - E(y_{t+J} | \Omega_{t-1}) \tag{13}$$

$$\Psi_{j,t}(H) = \frac{B_{H,t} \Sigma_t e_j}{\sqrt{\Sigma_{jj,t}}} \frac{\delta_{j,t}}{\sqrt{\Sigma_{jj,t}}}, \delta_{j,t} = \sqrt{\Sigma_{jj,t}} \tag{14}$$

$$\Psi_{j,t}(H) = \frac{B_{H,t}}{\sqrt{\Sigma_{jj,t}}} \Sigma_t e_j \tag{15}$$

e_j is an $m \times 1$ column vector with the m-th positional element being 1 and the remaining positional elements being 0 and *vice versa*.

Calculate $GFEVD\left(\tilde{\phi}_{ij,t}(H)\right)$, which represents the pairwise directional connectedness from j to i and accounts for the effect of the variable j on the variable i in terms of its forecast error variance share. These variance shares are normalised so that each row adds up to equal 1, which means that all variables together explain 100% of the prediction error variance of the variable. The calculations can be based on Equations 16–19:

(1) Total connectedness index

$$C_t(H) = \frac{\sum_{i,j=1,i\neq j}^{m} \tilde{\phi}_{ij,t}(H)}{\sum_{i,j=1}^{m} \tilde{\phi}_{ij,t}(H)} \times 100 = \frac{\sum_{i,j=1,i\neq j}^{m} \tilde{\phi}_{ij,t}(H)}{m} \times 100 \tag{16}$$

(2) Directional spillover index

Variable i from all other variables j:

$$C_{i\rightarrow j,t}(H) = \frac{\sum_{j=1,i\neq j}^{m} \tilde{\phi}_{ji,t}(H)}{\sum_{j=1}^{m} \tilde{\phi}_{ji,t}(H)} \times 100 \tag{17}$$

Variable i to all other variables j:

$$C_{t\leftarrow j,t}(H) = \frac{\sum_{j=1,i\neq j}^{m} \tilde{\phi}_{ij,t}(H)}{\sum_{i=1}^{m} \tilde{\phi}_{ji,t}(H)} \times 100 \tag{18}$$

(3) Net spillover index:

$$C_{i,t} = C_{i\rightarrow j,t}(H) - C_{i\leftarrow j,t}(H) \tag{19}$$

3.3 Frequency-domain spillover index

Based on Diebold and Yilmaz [29], and following Ding et al. [30], we construct a stationary VAR model as in Equation 20:

$$y_t = \Phi(L)y_t + \varepsilon_t = \sum_{i}^{p} \Phi_i(L)y_{t-i} + \varepsilon_t \tag{20}$$

where, y_t is the n-vector of returns of China's commodity and sectoral stock markets at time t. $\Phi(L)$ and ε_t presents the coefficient matrix and white-noise. Then, we give an infinite moving average form as in Equation 21:

$$y_t = \Psi(L)y_t + \varepsilon_t = \sum_{i}^{\infty} \Psi_i(L)y_{t-i} \tag{21}$$

where, Ψ_i is the infinite lag polynomial matrix when lag is i. Then, we calculate the GFEVD according to Equation 22:

$$\theta_{ij}^{g}(H) = \frac{\sigma_{jj}^{-1}\sum_{h=0}^{H-1}\left(\left(\Psi_h\Sigma\right)_{ij}\right)^2}{\sum_{h=0}^{H-1}\left(\Psi_h\Sigma\Psi_h'\right)_{ij}} \tag{22}$$

where, $\theta_{ij}^{g}(H)$ represents the contribution of the j-th variable to the variance of the prediction error of the i-th variable at horizon H. Σ is the variance matrix of errors matrix. σ_{jj} represents the j-th diagonal of Σ. Similar to the quantile-based spillover effect, we normalized the spillover effect of each market as Equation 23:

$$\tilde{\theta}_{ij}^{g}(H) = \frac{\theta_{ij}^{g}(H)}{\sum_{j=1}^{N}\theta_{ij}^{g}(H)} \tag{23}$$

According to Diebold and Yilmaz [29], the total spillover index can be calculated as Equation 24:

$$TSI(\tau) = \frac{\sum_{i=1}^{N}\sum_{j=1,i\neq j}^{N}\theta_{ij}^{-H}}{\sum_{i-1}^{N}\sum_{j=1}^{N}\omega_{ij}^{-H}} \times 100 \tag{24}$$

According to Baruník and Křehlík [31], we decompose the time domain impulse response Ψ_h to frequency domains by using Fourier transform as Equation 25:

$$\Psi\left(e^{-i\omega h}\right) = \sum_{0}^{\infty} e^{-i\omega h}\Psi_h \tag{25}$$

Then we calculate frequency domains of GFEVD as Equation 26:

$$\Theta_{ij}(\omega) = \frac{\sigma_{jj}^{-1}\left(\sum_{0}^{\infty}\Psi\left(e^{-i\omega h}\right)\Sigma\right)_{ij}^{2}}{\sum_{0}^{\infty}\Psi\left(e^{-i\omega h}\right)\Sigma\Psi\left(e^{-i\omega h}\right)_{ii}} \tag{26}$$

Then, we can obtain the pairwise spillover as Equation 27:

$$\tilde{\Theta}_{ij}(\omega) = \frac{\Theta_{ij}(\omega)}{\sum_{0}^{n}\Theta_{ij}(\omega)} \tag{27}$$

The pairwise connectedness at the frequency band (ω_1, ω_2) is calculated according to Equation 28:

$$\tilde{\Theta}_{ij}(\omega_1, \omega_2) = \int_{\omega_1}^{\omega_2}\tilde{\Theta}_{ij}(\omega)d\omega \tag{28}$$

We get the within net pairwise spillover between market i and market j as Equation 29:

$$C_{ij,net}^{d}(\omega_1, \omega_2) = \tilde{\Theta}_{ij}(\omega_1, \omega_2) - \tilde{\Theta}_{ji}(\omega_1, \omega_2) \tag{29}$$

The within total spillover on the band (ω_1, ω_2) is calculated according to Equation 30:

$$C^{(\omega_1,\omega_2)} = \frac{\sum_{i\neq j}^{n}\sum_{j=1}^{n}\tilde{\Theta}_{ij}(\omega_1, \omega_2)}{\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\Theta}_{ij}(\omega_1, \omega_2)} \tag{30}$$

The within from spillover, within to spillover and net spillover can be calculated according to Equations 31–33:

$$C_{i\leftarrow all}^{(\omega_1,\omega_2)} = \sum_{j\neq i}^{n}\tilde{\Theta}_{ij}(\omega_1, \omega_2) \tag{31}$$

$$C_{i\rightarrow all}^{(\omega_1,\omega_2)} = \sum_{j\neq i}^{n}\tilde{\Theta}_{ji}(\omega_1, \omega_2) \tag{32}$$

$$C_{i,net}^{(\omega_1,\omega_2)} = C_{i\rightarrow all}^{(\omega_1,\omega_2)} - C_{i\leftarrow all}^{(\omega_1,\omega_2)} \tag{33}$$

A spillover index essentially measures how volatility in one market "contagiously" spreads to other markets. For example, if the price of rare earths suddenly rises as a result of policy tightening, the spillover index is able to quantify the extent and duration of the impact of this event on the price of other metals. This is like a "risk map" for investors, identifying in advance which industries are likely to fluctuate in tandem; and for policymakers, it helps to assess the potential knock-on effects of policy interventions.

4 Data selection and descriptive statistical analysis

In this paper, we build on existing studies [23, 32–41], mapping key raw materials for each industry (Figure 1). Among the raw materials in the new energy industry, we focus on lithium, cobalt, nickel, copper, silicon and silver, which are mineral raw materials with strong co-production relationships, for example, cobalt seldom occurs in the form of independent minerals, and is mainly accompanied by iron, nickel, copper and other minerals; nickel is usually accompanied by minerals such as copper, iron, and platinum group metals, and the ores for the production of nickel are mainly copper-nickel sulphide ores at present; and one of the most important sources of silver is the production of copper and associated ores of lead-zinc ores, and silver is produced as a by-product of copper or lead-zinc production. Rare earth elements are symbiotic, and here we choose cerium, dysprosium, lanthanum, neodymium, terbium and praseodymium, which have wider application fields and larger consumption, as the core. The rapid development of the new energy industry has led to a sharp increase in the consumption of these metal raw materials, and the different raw materials show a clear co-consumption relationship, in which copper is widely used in wires, connectors and other components for conducting and transmitting electricity; lithium, cobalt and nickel are mainly used in battery-based energy storage devices; silver and silicon are the main materials for solar cells, which are used in the manufacture of photovoltaic cells in the main body structure; cerium, dysprosium, Cerium, dysprosium, lanthanum, neodymium, terbium, praseodymium and other rare earth elements are indispensable metal raw materials for permanent magnetic materials, hydrogen storage materials, luminescent materials and alloy materials.

As four resources, lithium, cobalt, nickel and silver, are dependent on large imports in China for their production, meaning that their international market prices have a significant impact on domestic prices, we have mainly used the global price indices provided by the International Monetary Fund when analysing these four markets. For the remaining eight raw materials and rare earth elements, Chinese domestic market prices are used (data from WIND database). This approach not only better reflects the impact of international price fluctuations of import-dependent resources on the domestic market, but also helps to understand the risk-transfer relationship between the domestic market and the international market, thus providing insights of more practical value for policymaking and market analysis. Based on data availability, the selected sample starts and ends from May 2014 to September 2024 and contains a total of 125 sample observations. The return series for each market are obtained through logarithmic differencing. The return series is the core series studied in this paper, and descriptive statistics are provided here, while the corresponding time series are plotted, i.e., shown in Figure 2. Table 1 shows the descriptive statistics, where skewness is used to measure the asymmetry of the data distribution. Positive skewness implies that the long tail is on the right; negative skewness is the opposite. Ce (cerium) shows a significant negative skewness of −3.24, indicating that the distribution is significantly skewed to the left. Kurtosis measures how sharp or flat the data distribution is. A high kurtosis means that it is sharper (has more extreme deviations) than the normal distribution. Si (silicon) and Nd (neodymium) show very high kurtosis of 30.50 and 29.73 respectively, implying that their data distributions have high peaks. In terms of kurtosis, the values are mostly greater than 3, indicating that the data distribution has sharp peaks and heavier tails, implying that there are more extremes (outliers) in the data, indicating a higher risk. The ADF values are the results of the unit root test, and the values show that all the series are smooth. The specific descriptive statistics values are shown by Table 1.

5 Empirical analysis

5.1 Total spillover

Figure 3 illustrates the overall spillover index between new energy-related metals and rare earth elements based on different order moment risk analyses, with higher values indicating a higher overall correlation between rare earths and raw material markets in the new energy sector. The shaded areas mark time periods with high risk spillovers due to policy impacts. The overall trend of the risk spillover between returns and volatility shows a more consistent upward trend, especially between 2016 and 2024. However, there are three key time periods in the total spillover index on volatility risk that are worth keeping an eye on - May 2014 to June 2015, March 2020 to November 2021, and July 2022 to September 2024. There are several policies in the three timeframes that have significantly impacted the market. For example, in May 2014, China released a draft Rare Earth Industry Development Plan (2016–2020), which aims to consolidate rare earth resources, improve the efficiency of resource utilisation, and stimulate domestic and foreign investment. Although this is a medium- to long-term plan, its expected effects caused market volatility in rare earth prices during the pre-implementation discussion phase, leading to investor sentiment swings and exacerbating risk spillovers.

In contrast, the number of time periods with high spillover of skewness risk is small and relatively short-lived, reflecting short-term irrational behaviour in the market, where investors may make quick decisions in the face of sudden good or bad news, leading to sharp price fluctuations in the market. However, such volatility is usually temporary and the market may quickly return to rationality as sentiment settles. Important periods of high spillover of such risks include May 2017 to May 2018 and June 2021 to March 2022. Among the typical policies are, in late December 2016, China's Ministry of Industry and Information Technology (MIIT) and eight other departments jointly issued a circular to carry out a special operation to crack down on rare earth violations nationwide from December 2016 to April 2017 in order to further regulate the order of the rare earth market. In July 2017, the Chinese government announced that it would continue to implement the rare earth export quota system and adjust the rare earth export quota for the second half of 2017 and the first half of 2018. In addition, the state conducted a total of three rare earth storage in 2017, and the storage price was gradually increased. These policies have intensified market concerns over rare earth supply, driving up prices in the international market and increasing the risk of bias in rare earth-related materials (e.g., lithium hydroxide, lanthanum, neodymium, etc.).

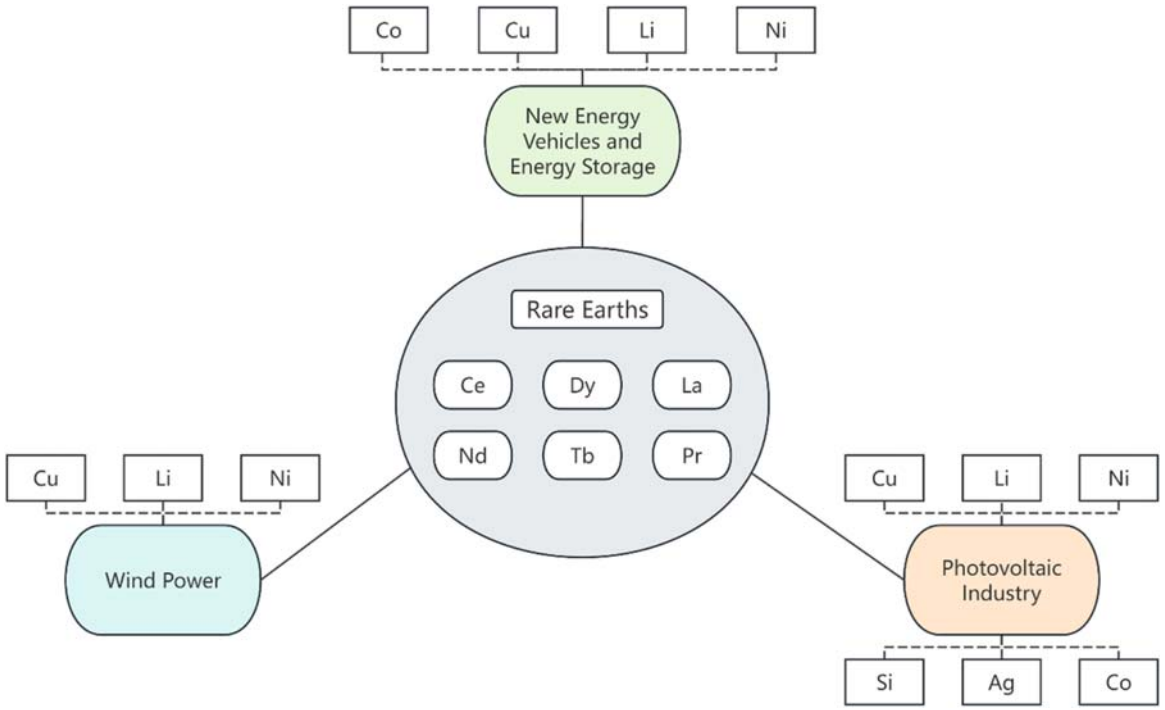


FIGURE 1
Critical raw materials by industry.



FIGURE 2
Time series of returns across markets.

TABLE 1 Descriptive statistics.

Metals markets	N	Max	Min	Mean	Std_Dev	Skewness	Kurtosis	ADF_Value
Li	125.00	−0.26	0.37	0.00	0.08	1.04	6.03	−7.20
Co	125.00	−0.31	0.23	0.00	0.09	−0.48	2.54	−3.60
Ni	125.00	−0.19	0.35	0.00	0.07	0.73	2.88	−4.89
Cu	125.00	−0.17	0.15	0.00	0.04	−0.34	3.37	−4.32
Si	125.00	−0.56	0.75	0.00	0.10	1.96	28.68	−5.14
Ag	125.00	−0.18	0.28	0.00	0.06	0.72	3.23	−5.47
Ce	125.00	−0.12	0.03	−0.01	0.02	−3.28	14.58	−4.71
Dy	125.00	−0.23	0.21	0.00	0.06	0.04	2.54	−6.49
La	125.00	−0.11	0.05	−0.01	0.02	−1.25	5.08	−3.58
Nd	125.00	−0.97	0.92	0.00	0.15	−0.40	27.88	−5.55
Tb	125.00	−0.25	0.21	0.00	0.08	−0.24	1.66	−5.29
Pr	125.00	−0.17	0.21	0.00	0.06	0.21	2.50	−4.39

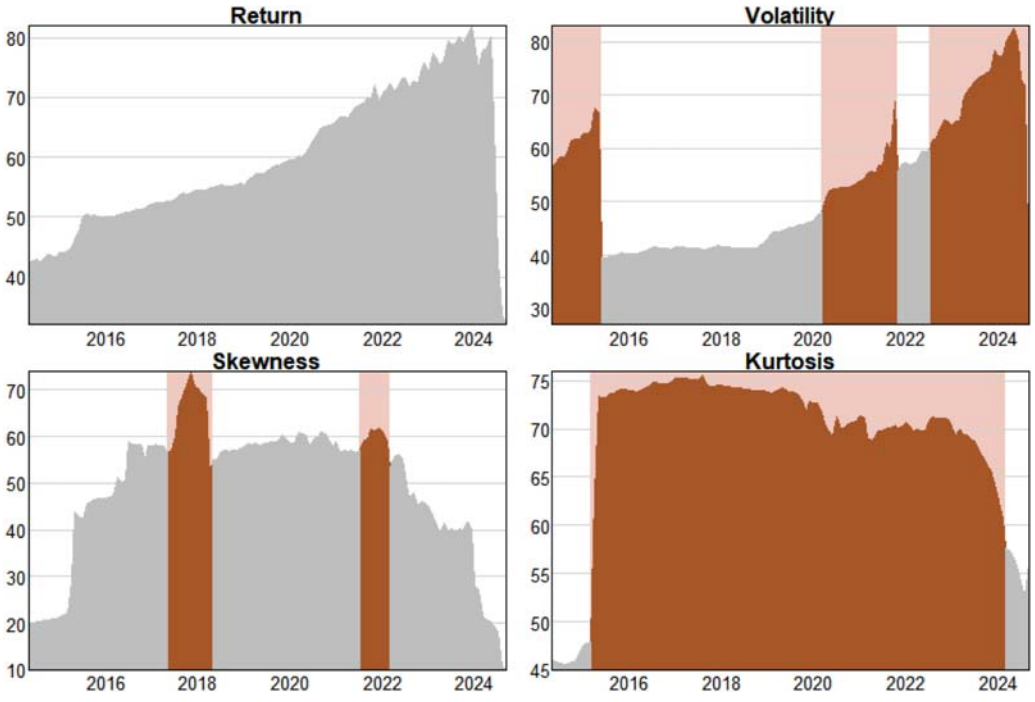


FIGURE 3
Total spillover index.

Cycles of high spillovers of kurtosis risk are longer in duration, reflecting the market's continued concern and patience with extreme situations or extreme returns. That is, when peak degree risk occurs in one market, investors may increase their alertness to potential crises, and this alertness can influence trading decisions in other markets, leading to the frequency of extreme volatility being correlated across markets. Specifically, between March 2015 and February 2024, spillover levels remain above 60% overall, suggesting

that on average, more than 60% of the kurtosis risk shocks to each variable come from other variables.

Combining the timing of high spillovers for different risk types, we are particularly concerned about the period from June 2021 to November 2021, when all four types of risks show a significant uptick. Several factors drove this scenario. First, government policies supported the rapid development of the new energy industry. Second, controls on rare earth resources tightened. Third, geopolitical competition for resources intensified. For example, in 2021 China implemented regulatory measures on rare earth production. These aimed to reduce overcapacity and improve environmental protection.

According to the theory of industrial organisation, policy adjustments will directly affect the cost structure and competitive landscape of enterprises. The rare earth production restriction policy has compressed supply and pushed up raw material prices in the short term, but due to the lack of clear market expectations about the future direction of the policy, it is difficult for enterprises to form stable long-term investment decisions. This divergence in expectations has led to sharp fluctuations in investor sentiment, with prices deviating from fundamentals and triggering an overreaction in the market. In addition, as the rare earth industry has a high degree of concentration and barriers to entry, policy changes will affect downstream new energy companies through the supply chain step by step. With asymmetric information, downstream companies may stockpile inventory or adjust their purchasing strategies due to concerns about supply shortages, further exacerbating price volatility and market instability. This process fully reflects the complexity of market dynamics under changing policy environments, and confirms the logic of industrial organisation theory about the interaction of policy shocks, market structure and irrational behaviour.

5.2 Net spillover

The NET Spillover Index shows information about the direction of volatility, with a negative (positive) value meaning that a market receives (transmits) more risk spillover than it transmits (receives). Markets that show significant positive values are the main risk emitters within that system. The main risk premium emitters in the four types of risk premium systems mostly contain rare earth element markets, as described below:

Figure 4 shows a graph of net return risk spillovers. Terbium has long been the main risk emitter within the system in this type of risk. Silicon and lanthanum, on the other hand, have positive net spillover indexes until 2022, but have shifted to negative as the main risk takers in recent years. This shift may be due to multiple factors such as industry dynamics, market demand fluctuations, supply chain issues and changes in the policy environment, which have led both to show greater vulnerability to external shocks and increased sensitivity to market volatility. In contrast to silicon and lanthanum, dysprosium has a predominantly negative net spillover index until 2023, but has become an external emitter of risk in recent years. In contrast, only cobalt and cerium have had negative net spillover indexes for a long period of time, and both are highly vulnerable to fluctuations in return risk in other markets.

The net volatility spillover (Figure 5) shows cobalt, dysprosium and neodymium as the main risk emitters from May 2014 to June 2015, with neodymium in particular showing extremely significant risk spillover during this period. In early to mid-2015, China implemented a series of regulatory policies targeting rare earth resources. By restricting rare earth mining and export quotas, China increased its management of this important resource. These policies triggered significant changes in supply and demand, with neodymium in particular attracting significant attention due to its key applications in new energy sectors such as NdFeB permanent magnets used in electric vehicles and wind turbines. As uncertainty about the market supply of neodymium intensified, companies and investors reacted strongly to price fluctuations in neodymium, leading to spillover of volatility risk to the new energy materials market and other rare earth element markets.

During the period from March 2020 to November 2021, silicon and lanthanum become the main volatility risk emitters. With the development of new energy industries such as photovoltaics, silicon as the main material for photovoltaic cells, and lanthanum with its high light-absorption capacity and wide bandgap characteristics, which can significantly improve the photovoltaic conversion efficiency, stability, and lifespan of solar cells, the co-consumption of silicon and lanthanum is becoming more pronounced, and the complementary nature of the two improves the overall efficiency of the photovoltaic system. Other new energy materials and rare earth elements are dependent on the performance of silicon and lanthanum in the supply chain, which makes them risk takers, and when silicon and lanthanum price volatility intensifies, the overall supply chain and price stability are impacted, creating a new market risk transmission mechanism.

After July 2022, a new situation emerged in the market. Copper, silver and terbium have become major volatility risk emitters, a change largely driven by the global economic recovery, green technology transition and accelerating growth in the electric vehicle industry. As renewable energy sources and electric vehicles gain popularity, the importance of copper as a key conductor in batteries, charging infrastructures and other electronic devices has risen significantly, leading to a surge in its demand. Meanwhile, increasing use of silver in photovoltaic cells and highly conductive materials has also contributed to its market volatility. Moreover, rising consumer demand for terbium, a rare earth material widely used in high-efficiency motors and green technology devices, further pushed up prices and volatility. A series of macro policies, such as national support policies for green energy and electric mobility, have increased the demand for these materials, causing their supply and demand to change, with a corresponding increase in volatility and risk, creating new market headwinds.

Skewness and kurtosis, as quantifications of extreme risk spillovers, have similar spillover patterns (See Figures 6, 7). Neodymium is the main risk emitter of skewness and kurtosis risk. This is mainly due to the high consumption and volatility of supply and demand of neodymium[2], which directly affects the production cost and market price of related materials (e.g., NdFeB permanent magnets contain 25%–32% of rare-earth elements, of

2 Details can be found in the survey report Critical Minerals Datase. https://www.iea.org/data-and-statistics/data-product/critical-minerals-datase.

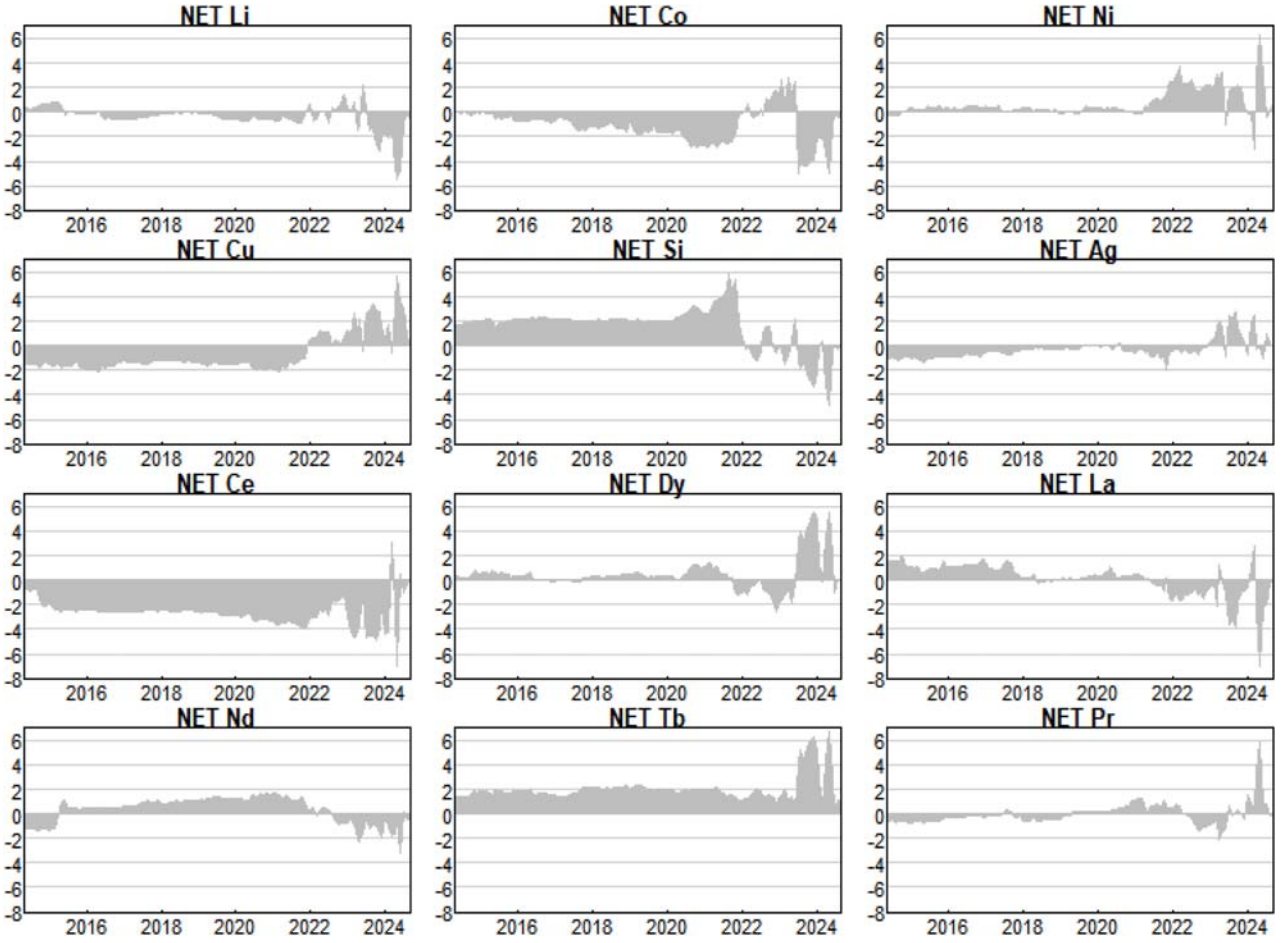


FIGURE 4
Net spillover of return.

which neodymium accounts for more than 90%) [42]. As an important rare earth element, the mining and supply of neodymium is often affected by multiple factors such as policy, environment and international trade. Meanwhile, due to the irreplaceable nature of neodymium in permanent magnet materials, its supply elasticity is low, resulting in price fluctuations that are more easily transmitted to downstream industries, further exacerbating its volatility in the market. When the market price of neodymium fluctuates dramatically, it not only triggers investors in the new energy materials market to react quickly and adjust their investment strategies, but also causes a spreading effect of kurtosis and skewness risk in other related markets due to concerns about rising costs and supply uncertainty, thus exacerbating volatility across the market.

Although cerium, dysprosium, lanthanum, terbium and praseodymium are also used in new energy materials, their market demand is relatively stable and their price movements are small, lacking the dramatic volatility similar to neodymium. Cerium is mainly used in catalysts, dysprosium is used in high-temperature superconducting materials, while lanthanum, terbium and praseodymium are also mostly used in relatively niche applications, where the demand is not as concentrated and strong as neodymium's, and thus its influence is limited and fails to create significant kurtosis risk. This difference makes neodymium more prominent as a risk emitter in the market for new energy materials.

5.3 Static spillover analysis

For static return spillover (see Table 2), the FROM column shows the sum of accepted return risk spillover. In this column, the maximum values for new energy-related metals and rare earth elements are 63.8% and 71%t for copper and terbium, respectively. They are the main recipients of return risk in their respective areas. Correspondingly, copper accepts risk mainly from nickel and silver, and terbium mainly from praseodymium and neodymium. The TO row is the sum of the market's share of risk spillovers to other objects, and in this row we focus on nickel and dysprosium, which are the main senders of volatility and to which markets should pay extra attention for the impact of spillover yield risk. The NET row shows the difference between the sum of the columns and the sum of their own rows, which provides the share of net volatility shocks, with a maximum value of 18.6 percent for silicon and a minimum value of −31.5 percent for cerium.

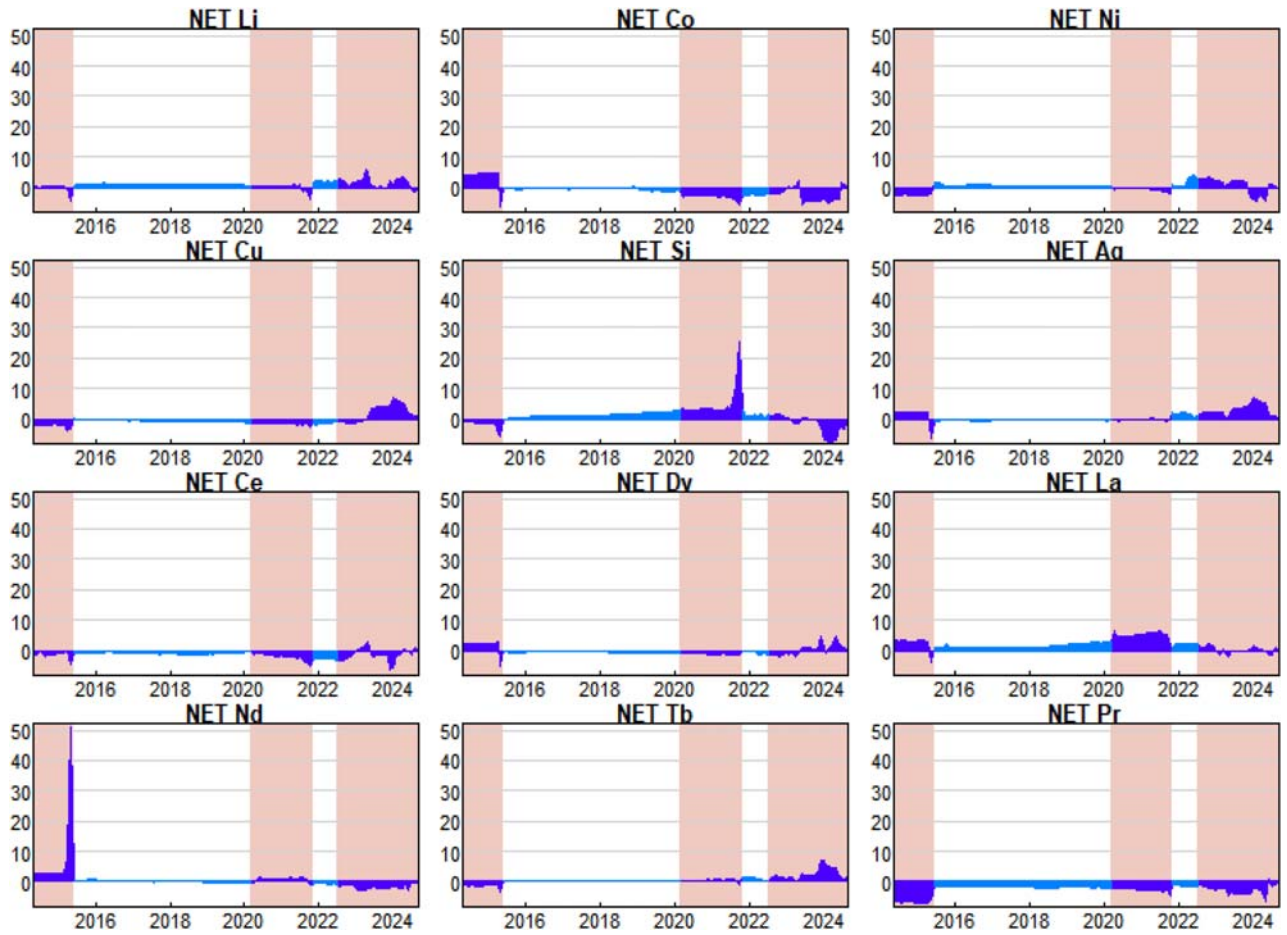


FIGURE 5
Net spillover of volatility.

In terms of volatility (see Table 3), some of the values in the "FROM" column are slightly above returns, represented by silver and terbium, which receive the most volatility risks. Silver is particularly exposed to the volatility of copper and terbium, which we consider to be due to the co-consumption and co-production of silver, copper and terbium, mainly in electronics and high-tech applications. Silver and copper are widely used in electrical conductors, while terbium, a rare earth element, often plays an important role in electronic devices such as fluorescent displays and computer memory components. The smelting of copper ores is often accompanied by the extraction of silver in the production process. In addition, market dynamics of copper and terbium can influence the demand for silver, e.g., fluctuations in copper prices can lead to changes in the cost of electronics, which in turn affects silver consumption. Thus, the interrelationship between the three not only promotes the integrated utilisation of resources, but also creates a strong link at the technical and market levels. The volatility risk of terbium, on the other hand, arises mainly from lithium and dysprosium, which play an important role in the new energy and high-tech industries, with lithium being a key component of lithium-ion batteries and dysprosium being an important rare earth element used in the manufacture of high-performance permanent magnet materials.

Compared to the first two types of static spillovers, the static spillovers of skewness and kurtosis are a more intuitive reflection of the dominant role of neodymium in extreme risk spillovers, with significantly higher values of external spillovers than in other markets (see Tables 4, 5). In skewness risk, neodymium mainly affects nickel and terbium. In the kurtosis risk scenario, neodymium mainly affects cobalt and praseodymium. The impact of neodymium on nickel and terbium in the skewness risk spillover scenario reflects the asymmetric response of the market to the technological applications of these metals. Nickel, as an important component of battery materials and alloys, is susceptible to fluctuations in neodymium supply, leading to skewed price movements, while terbium's importance in photovoltaic applications makes it highly sensitive to changes in neodymium supply and demand, which in turn leads to skewness risk spillovers. In comparison, kurtosis risk spillover focuses on the fatness of the tails of the return distribution, indicating a higher probability of extreme events. The impact of neodymium is more pronounced for cobalt and praseodymium, with cobalt having a strong correlation with neodymium market dynamics in the context of high demand

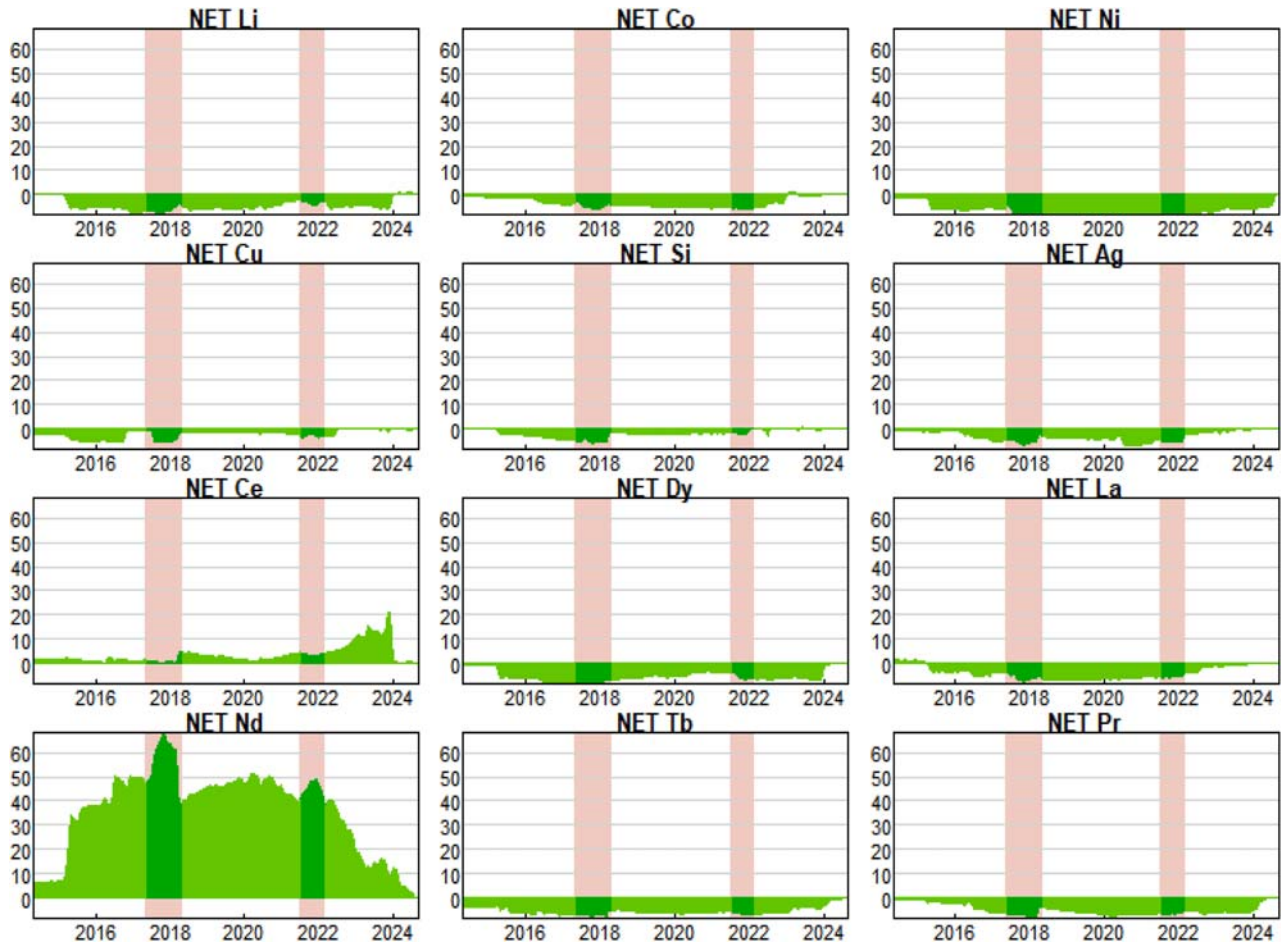


FIGURE 6
Net spillover of skewness.

scenarios from the electric vehicle industry and new energy industry. Praseodymium and neodymium are both light rare earth elements with a close co-production relationship, and the rapid increase in neodymium consumption has created a surplus in the praseodymium market [26]. When the supply of neodymium fluctuates, especially in the event of production disruptions or policy adjustments, the supply and price of cobalt and praseodymium are also affected, leading to extreme price volatility in their markets, reflecting a clear kurtosis risk spillover effect.

5.4 Evolution of the maximum pairwise spillover index

As mentioned earlier, co-consumption and co-production are common and important material dependencies in industrial production and economic theory. Co-production can affect synergistic price fluctuations through the cost transmission mechanism of the chain. Co-consumption, on the other hand, generates price synergies through end-product demand linkages, but is more resilient than co-production, as it is more subject to technological substitution and market regulation. The Maximum Pairwise Spillover Index can reveal the interplay of these materials in market supply and demand changes, and identify which bilateral relationships are most likely to be potentially correlated. The time-varying characteristics of strongly correlated markets in return, volatility, skewness and kurtosis risk spillovers can also be assessed visually through dynamic evolutionary analysis. In Figure 8, the right axis represents the different pairs of metals, and the left panel shows the corresponding maximum risk premium values.

In return risk spillover, the largest source of risk before 2017 showed a strong randomness, mainly due to the degree of dependence between markets has not yet formed a stable structure, which makes the interactions and spillover effects between different metals more complex, silver-lithium, silver-nickel and other metal combinations have once become the largest paired spillover index. However, since 2017, with the advancement of technology, changes in market demand, rapid development of the new energy industry and the gradual maturation of the industrial chain, specific metal combinations such as lanthanum-cobalt, neodymium-cobalt and dysprosium-cerium have often become strong correlation combinations, demonstrating more stable interrelationships, which reflects the new energy market's preference for these specific material combinations in recent years.

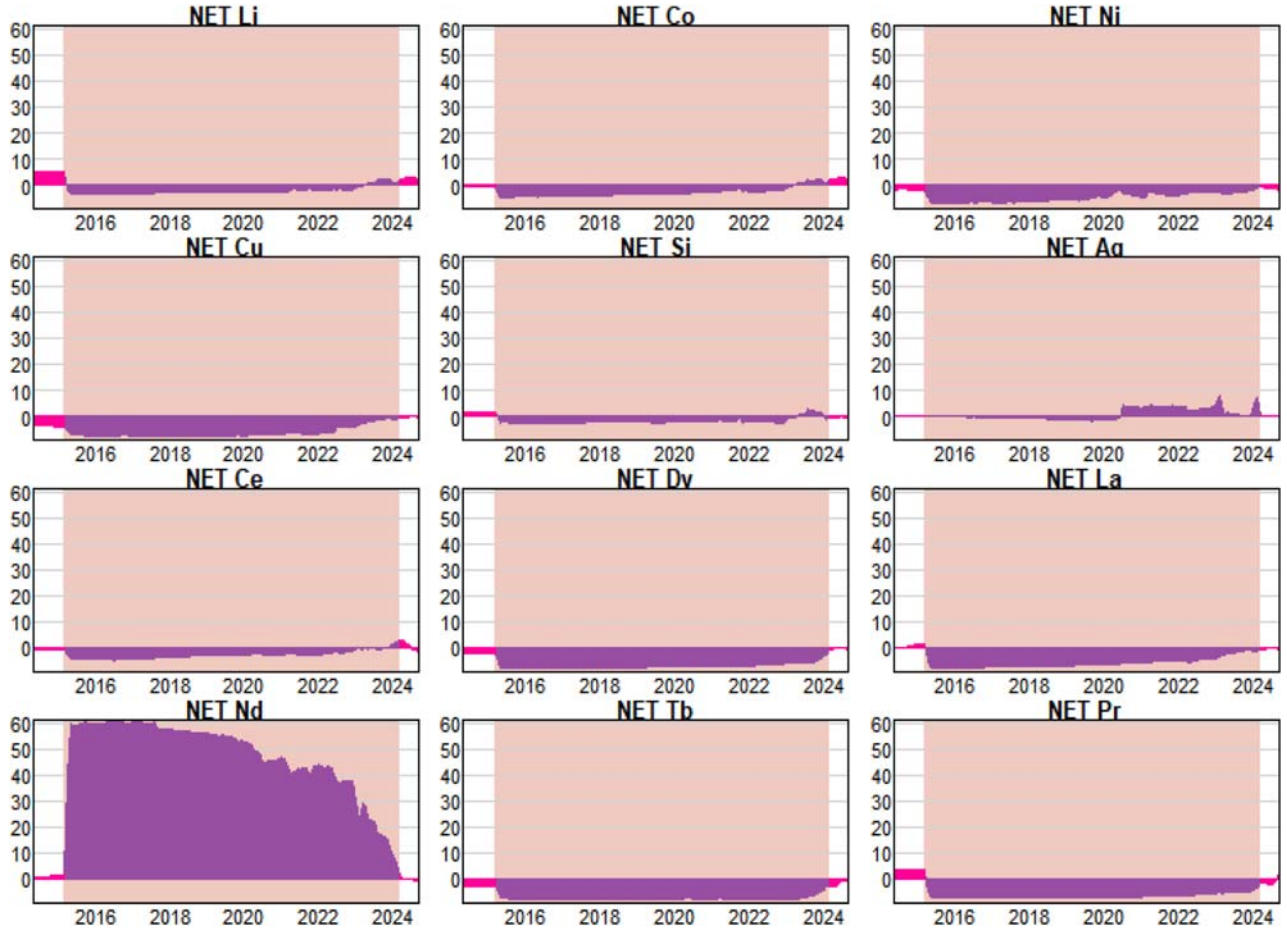


FIGURE 7
Net spillover of kurtosis.

Specifically, all three pairs of metals have significant co-consumption or co-production relationships. In certain specific alloys and battery materials, lanthanum and cobalt are often used together to enhance material performance and functionality. Lanthanum improves the electrochemical properties of nickel-metal hydride (NiMH) batteries and other materials, while cobalt plays a key role in batteries, boosting energy density and cycle life. Market demand for both is therefore closely linked, highlighting their synergistic effects. Similarly, neodymium, a rare earth element, is often used in combination with cobalt to improve the magnetic properties and thermal stability of permanent magnet materials. This synergy makes neodymium and cobalt closely linked in the market, leading to a symbiotic economic relationship. Dysprosium and cerium, as rare earth elements, have a clear co-production relationship, and they tend to show a high degree of interconnectedness in sales and market demand, which leads to a relative concentration of risk spillovers in returns.

In volatility risk, the value of the maximum risk premium is relatively stable until the end of 2023, when higher values occur, arising between neodymium-silver. The two also have a co-consumption relationship. For environmental protection and energy saving considerations, most countries now have great support for new energy vehicles in terms of policy, which has shifted cars from internal combustion engine dominance to new energy sources. In the application of electric vehicles, neodymium as a rare earth element is used to manufacture high-performance motors to enhance the efficiency of the power system. Silver is mainly used in automobiles for various types of switches, circuit boards, circuit breakers and brazing alloys due to its superb conductivity, while electric vehicles tend to be more highly electronic, so the demand for silver has increased significantly. As a result, the price of neodymium, a key rare earth element, has increased in volatility due to soaring demand. At the same time, the importance of silver in batteries and electronic products has led to policy support and market interest. The combination of their applications in the electric vehicle industry sector has led to a sharp increase in risk spillover between neodymium and silver. The market connection between neodymium and silver becomes tighter as the policy effect is quickly apparent, but as the market adapts to the policy impact and strengthens its risk management, the spillover between the two will also experience a transient decline, reflecting the dynamic balance between market behaviour and policy impact. Thus, the risk spillover between neodymium and silver is characterised by both

TABLE 2 Static spillovers of return.

Metals markets	Li	Co	Ni	Cu	Si	Ag	Ce	Dy	La	Nd	Tb	Pr	From
Li	40	2.8	8.4	5.4	1.9	6.7	6.6	3.2	4.3	7.1	7.1	6.5	60
Co	5.3	42.4	4.1	1.8	4.4	4.4	1.7	5.9	8.5	8	6.6	6.8	57.6
Ni	6.9	5.2	38.1	13.4	1.8	10.7	3	3.1	4.2	2.7	7	4	61.9
Cu	3.6	6.2	12.1	36.2	3.9	11.9	2.2	4.1	6.8	2.6	7.6	2.8	63.8
Si	2.3	4.4	1.4	2.3	70.3	4.1	1.7	2.7	4.8	1.4	1.9	2.8	29.7
Ag	2.2	4.6	12.5	11	9	42.1	2	5	3.1	1.7	5.5	1.4	57.9
Ce	9.8	4.8	5.1	6.4	2.3	3.9	36.3	8.7	4.5	6.2	7	5	63.7
Dy	4.1	1.2	4.7	4.4	5.6	4.6	2.3	32.6	2.5	10.2	21	6.9	67.4
La	4.7	3.6	2.6	2.8	6.3	3.1	4.8	4.5	52.1	5.2	3.5	6.7	47.9
Nd	5.8	3.3	5.1	1.6	4.7	1.4	2.8	10.1	4.1	31.6	15.4	14.2	68.4
Tb	4.7	4.6	6.7	4.3	4.9	2.5	1.8	17.3	2.7	12.1	29	9.3	71
Pr	5.8	3.8	7.6	2.1	3.5	1.3	3.3	8.1	3.7	15	12.6	33.1	66.9
TO	55.3	44.6	70.2	55.5	48.3	54.5	32.2	72.9	49.2	72.2	95	66.3	TCI
NET	−4.7	−13	8.4	−8.4	18.6	−3.4	−31.5	5.4	1.2	3.8	24	−0.5	59.7

TABLE 3 Static spillovers of volatility.

Metals markets	Li	Co	Ni	Cu	Si	Ag	Ce	Dy	La	Nd	Tb	Pr	From
Li	47.6	4.9	4.1	2.4	6.5	3.8	9.1	3.5	7.7	4.5	3.2	2.6	52.4
Co	6.5	42.2	4.9	3.1	4.4	6.4	6.9	3	12.1	5.4	4	1.1	57.8
Ni	3.2	4.3	53.6	3.5	4.4	2.3	6.2	3.7	6	3.3	3.6	5.9	46.4
Cu	2.1	2.1	1	51.8	1.5	15.5	1.7	3.1	2.5	1.4	16.3	1	48.2
Si	3.1	2.3	1.4	1.8	57.2	8.6	2.3	2.1	7.1	5.7	7.5	1	42.8
Ag	4.1	6.5	4.1	10.8	4.9	37.6	2.5	2.8	4.9	7.6	11.1	3.2	62.4
Ce	7.1	6.2	13.1	3.8	10.7	4.5	37	3.3	6.2	1.9	3.4	2.8	63
Dy	5.8	3.4	5	1.1	2.7	3.2	3.1	53	2.6	3.6	13.2	3.5	47
La	5	3.3	2.8	3.4	4.7	8.2	2.6	2	55.9	7.5	3.8	0.8	44.1
Nd	4.8	2.7	5.3	1.5	2.7	3.1	4	2.8	7.2	55	5.1	5.8	45
Tb	13.7	2.1	3.2	9.6	5.5	11.5	1.9	11.5	4.1	3.2	32.2	1.7	67.8
Pr	8.1	8.1	3.5	1.9	5.6	3.6	6.2	5.6	5.6	7	5.1	39.7	60.3
TO	63.8	45.8	48.4	42.8	53.6	70.7	46.5	43.3	66	50.9	76.1	29.5	TCI
NET	11.4	−12	2	−5.4	10.7	8.3	−16.6	−3.7	21.9	5.9	8.3	−30.8	53.1

10.3389/fphy.2025.1616099

TABLE 4 Static spillovers in skewness.

Metals markets	Li	Co	Ni	Cu	Si	Ag	Ce	Dy	La	Nd	Tb	Pr	From
Li	30	0.6	0.3	0.7	0.4	0.6	11	4.2	3.1	46.9	1.4	0.8	70
Co	0.8	57.4	0.7	1.8	0.7	0.5	1.5	0.5	1	34.1	0.6	0.3	42.6
Ni	3.3	0.3	19.9	1.3	0.4	1.2	1.2	0.5	1.5	66.8	3.1	0.5	80.1
Cu	1.5	0.9	1.1	66.3	8.1	0.6	0.9	0.8	0.4	18.8	0.5	0.2	33.7
Si	0.3	0.3	0.2	0.9	63.5	0.6	0.6	0.3	2.1	27.8	0.6	2.7	36.5
Ag	2.7	0.4	0.8	0.6	0.7	55.8	1.9	0.7	1.8	32.3	1.5	0.7	44.2
Ce	0.1	0	0.3	0	0	0.2	98.2	0.1	0	0.8	0.1	0.2	1.8
Dy	5.6	0.8	0.2	0.8	1.4	1.2	5.6	27.7	0.5	51	4	1.1	72.3
La	0.5	0.3	0.6	1	2.8	0.7	0.6	0.7	42.7	48.9	0.6	0.7	57.3
Nd	0.1	0	0	1.3	0	0	0	0	0	97	1.3	0.2	3
Tb	1.1	0.3	1	0.7	0.4	1.1	13.7	1.9	0.8	57	19.5	2.5	80.5
Pr	1.5	0.9	0.4	0.9	1.5	3	6.9	2.4	1.6	45.2	2.8	33	67
TO	17.7	4.9	5.7	9.9	16.4	9.8	43.9	12	12.9	429.7	16.3	9.8	TCI
NET	−52.3	−37.7	−74.4	−23.8	−20.1	−34.3	42.1	−60.2	−44.4	426.7	−64.2	−57.2	49.1

TABLE 5 Static spillovers in kurtosis.

Metals markets	Li	Co	Ni	Cu	Si	Ag	Ce	Dy	La	Nd	Tb	Pr	From
Li	14	11	1	2.6	6.4	1.1	2.7	2.4	6.3	47.3	0.4	4.7	86
Co	5.6	7	0.5	1.3	2.8	0.5	0.9	1.2	3	74.2	0.2	2.9	93
Ni	4.9	3.5	26.6	4.4	3	1.8	8.9	2	1.9	39.5	0.4	3	73.4
Cu	4.4	3.8	1.3	11.1	2.9	0.5	0.8	1	3.9	67.9	0.7	1.8	88.9
Si	18.9	15.9	1.7	2.9	43.5	0	3.7	1.9	6.4	2.1	0.1	2.9	56.5
Ag	0.8	0.7	0.1	0.2	0.3	93.7	0.2	0.2	0.2	3.5	0.1	0.2	6.3
Ce	9.4	7.4	12.9	1.2	3.8	0.1	48.1	1.9	2.1	8	0	5.2	51.9
Dy	4.8	4.3	1	1.4	2.8	4.9	1.5	9.5	2.6	62.7	1.3	3.3	90.5
La	5.8	5	0.6	1.5	2.5	0.5	1.1	1.1	7.7	70.7	0.5	3	92.3
Nd	1.8	1.6	0	0.3	0.9	0	0.1	0.2	0.5	94.4	0	0.3	5.6
Tb	6.5	5.6	0.6	1.3	8.8	3.7	1.7	2.4	2.4	51.8	13.5	1.7	86.5
Pr	6.1	5.2	0.7	1.2	2.9	0.5	1.2	1.8	3	71.3	0.3	5.7	94.3
TO	69	63.9	20.4	18.3	37.1	13.6	22.9	16.2	32	498.8	3.9	28.9	TCI
NET	−17	−29.1	−53	−70.6	−19.3	7.3	−29	−74.4	−60.2	493.2	−82.6	−65.4	68.8

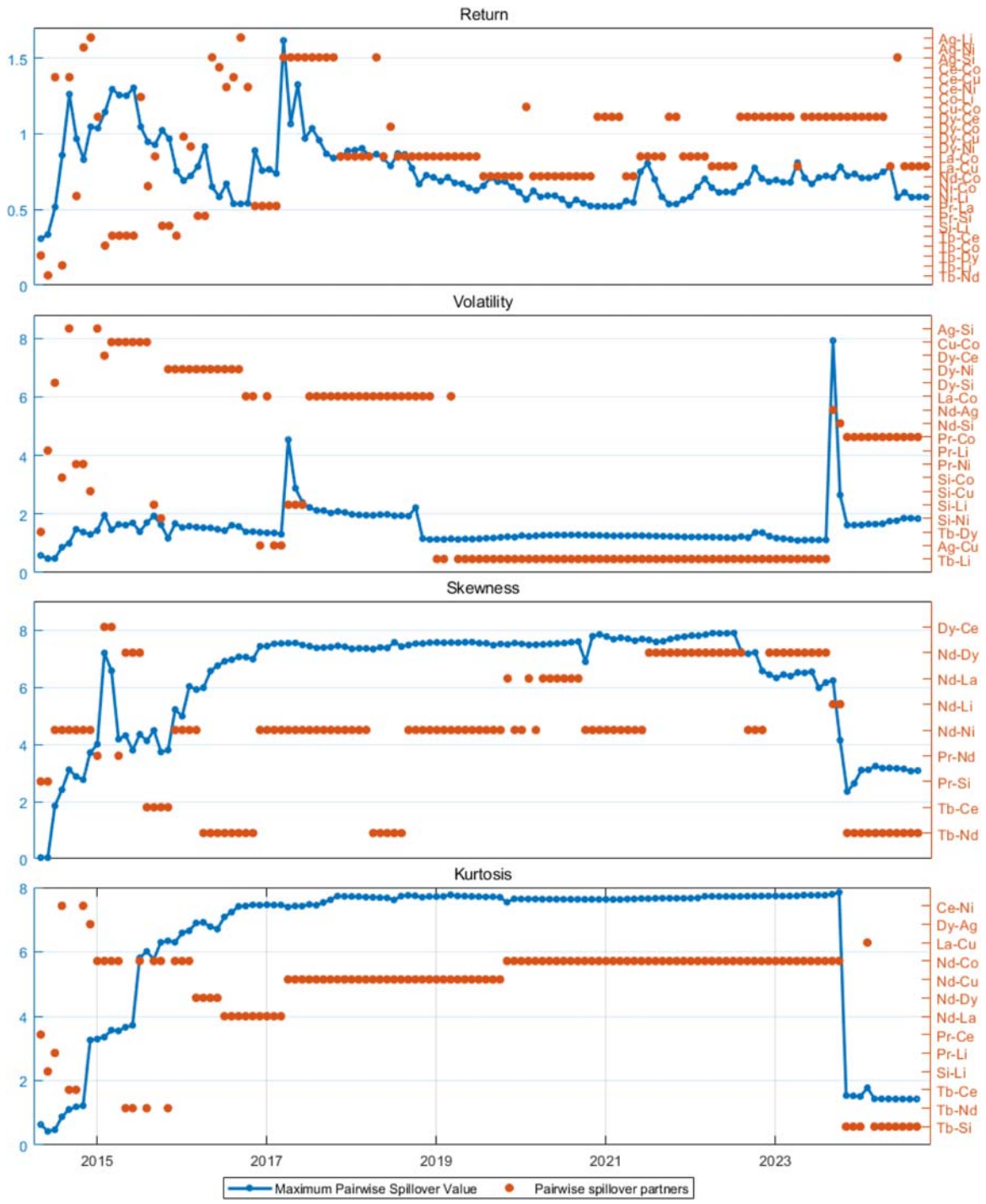


FIGURE 8
Maximum pairwise spillage evolution.

policy advances and transient fluctuations due to adaptive changes in supply and demand.

In the skewness risk premium, the value of the maximum risk premium is significantly higher than the return and volatility. This suggests that there are strong asymmetric price volatility risk-transfer effects in the market. In particular, in extreme cases, risk linkages between certain assets may increase significantly. This suggests that investors need to pay special attention to skewness risk when managing risk. And, as can be seen from the right axis, these maximum risk spillovers are more concentrated in a few risk sources, namely: neodymium-dysprosium and neodymium-terbium. Neodymium and dysprosium are commonly used together in electric vehicle motors, and they are interdependent on each other, affecting the performance and cost of the common



FIGURE 9
Evolution of the spillover network (return (left), skewness (centre), kurtosis (right)).

application. These two rare earth metals have both a co-production relationship and a co-consumption relationship in specific industries, highlighting their correlation and joint influence on extreme risk volatility. The high spillover risk between neodymium and terbium, on the other hand, considers their co-production relationship. The existence of this relationship is significant in terms of economics and resource use, as neodymium can be extracted while terbium can be recycled more efficiently, maximizing resource use.

Among the kurtosis risk spillovers, the largest pairwise risk spillovers have a more concentrated source, mainly from neodymium-cobalt and neodymium-copper. The neodymium-cobalt relationship has been mentioned before, and it is important to note here that the result confirms the high degree of connectivity between the two markets in terms of the potential for abnormal volatility in the face of external shocks. As for neodymium-copper, there is a co-consumption relationship between the two, especially in technology areas such as electric motors and batteries. Neodymium, an important rare earth element, is used to make high-performance permanent magnet materials, while copper is widely used in electrical connections and conductive components. Together, these two materials play a key role in high-tech products such as electric vehicles, so when it comes to design and material selection, manufacturers need to consider both neodymium and copper in terms of performance and cost in order to optimise overall performance. In this context, it is particularly important to focus on the risk transfer of extreme price fluctuations between copper and neodymium.

As a result, from the perspective of industrial linkages, it can be seen that the intensity of risk transmission is more pronounced in metal combinations of co-consumption and co-production, owing to their key position in the industrial chain, technological lock-in effect and high industrial concentration. Such portfolios are subject to fixed ratios and lack elasticity of substitution, with fluctuations in supply and demand spreading rapidly through production formulas and procurement networks, creating a chain reaction, whereas portfolios with strong substitutability have a weaker risk transmission.

5.5 Spillover network evolution

Based on the paired spillover index, the spillover network is plotted here for the high spillover period to visualise the directionality and complexity of the risk spillovers. The period from June 2021 to November 2021 corresponds to the time of analysis (based on the previous analysis). Since the evolution of the spillover network for return, skewness, and kurtosis is not obvious, the static spillover network is plotted here for the analysis (See Figure 9). The grey network is return, the green network is skewness, and the pink network is kurtosis. During this time period, the markets with strong external risk spillovers continued to be dominated by rare earth elements.

For the return spillover network, silicon and lanthanum are the main spillover emitters, and silver and cobalt are the corresponding risk takers. Cerium is the central return risk taker of all the elements, and most of the metals' return risks are passed on to it. The most important spillover emitter in the skewness spillover network is neodymium, which can be seen to transmit a large amount of kurtosis risk to nickel, praseodymium, terbium, cobalt and lanthanum. The kurtosis network bears some similarity to the skewness network, with neodymium remaining dominant in the kurtosis risk network. The biggest difference is the shift in terbium's position. It has some ability to emit risk in both the return network and the skewness network, but in the kurtosis network it has become a core risk taker and is highly susceptible to risk shocks from other markets. It follows that in the skewness risk spillover network, it is able to influence the risk dynamics of other markets by transmitting asymmetric risk. However, in the kurtosis risk spillover network, it is highly vulnerable to high-frequency extreme volatility due to its high sensitivity to extreme market fluctuations.

The blue network is a volatility spillover network with a clearer evolution (See Figure 10). As can be seen from the colour shades of the edges in the network, spillovers in the system increased overall during this period, mainly due to silicon and lanthanum. Both strengthened their external risk spillovers, targeting markets such as cerium, terbium, cobalt, praseodymium and lithium. In contrast, the metals markets associated with the

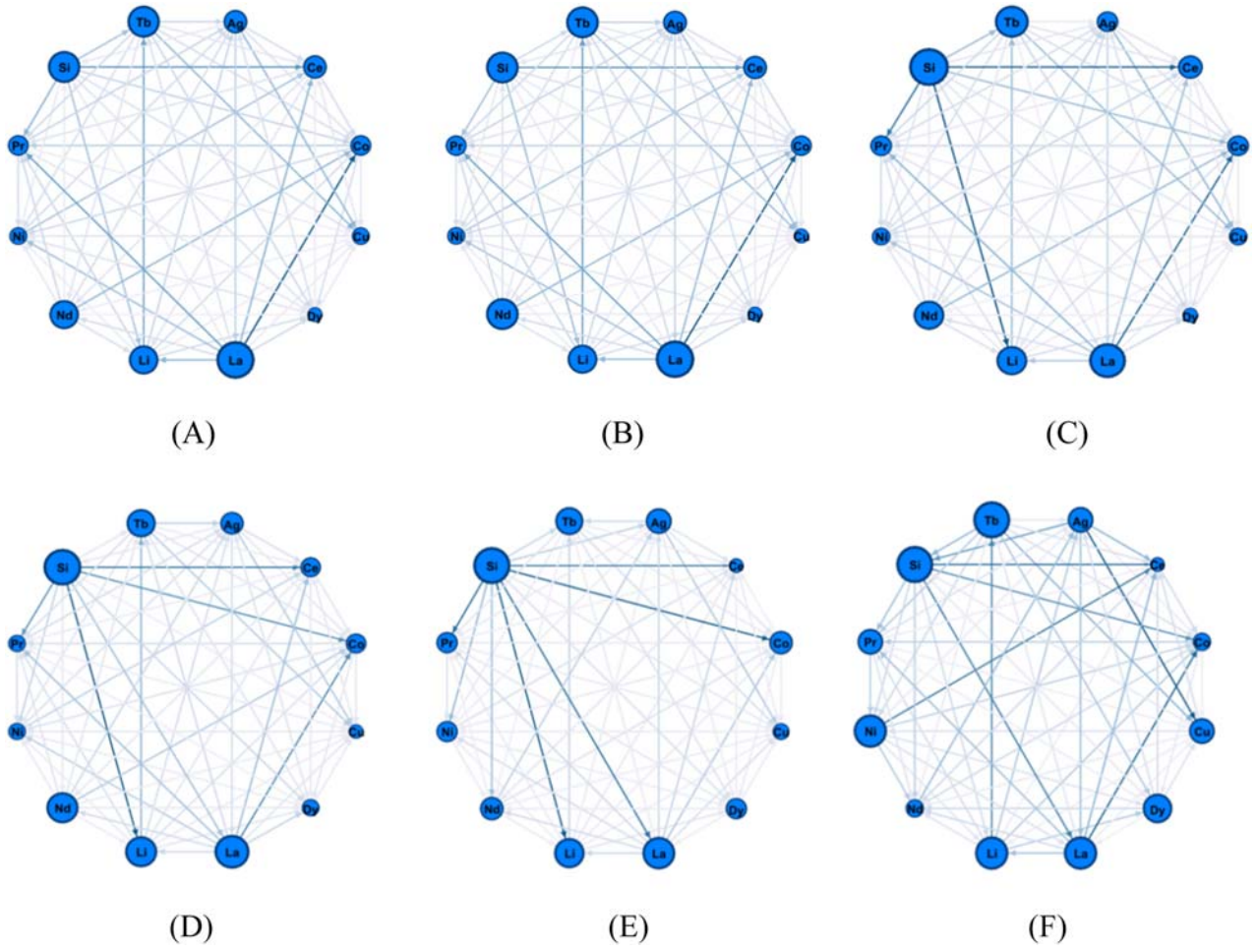


FIGURE 10
Evolution of the volatility spillover network ((**A–F**)represent June to November 2021)

new energy sector are more likely to evolve as risk receivers. The network bilateral relationship between most raw materials and rare earths markets is one in which the latter points to the former. This situation highlights the key position of rare earths in the new energy industry chain. Rare earth price fluctuations have a direct impact on the cost and competitiveness of new energy products, and all parties in the industry need to pay more attention to the volatility risks of the rare earth market in order to seek effective risk management and response strategies.

5.6 Comparative analysis of long and short-term market risk spillover

Here, we perform a comparative analysis of long- and short-term market spillovers for each order of momentary risk (See Figures 11–14 for visualisation). The Chinese rare earth market has a more pronounced pattern of long- and short-term correlation. The outward spillovers of different order moments generally show significant positive long- and short-term correlations. In the case of return risk spillovers, for example, when cerium and terbium have a significant impact on other markets in the short run, they also have a significant impact on other markets in the long run. This reflects the fact that uncertainty and shock volatility in rare earth markets in the short term can carry over and build up in the long term, triggering broader changes in market sentiment and asset price adjustments. Generally speaking, the impact of risk transmission in the long term is more far-reaching and may lead to adjustments in the overall market structure and asset allocation, as well as affecting macroeconomic policies and the direction of long-term market development. Therefore, the impact of risk volatility in rare earth markets on other markets in the short term will manifest itself in more significant and lasting effects in the long term.

It is worth mentioning that when analysing the time-domain spillover index, we obtained an important finding. Neodymium exhibits a strong outward influence in terms of skewness and kurtosis spillovers. This may be attributed to the model's ability to capture market dynamics and volatility characteristics in real time. As a result, neodymium is able to quickly show its impact on market sentiment and risk transmission. In contrast, when

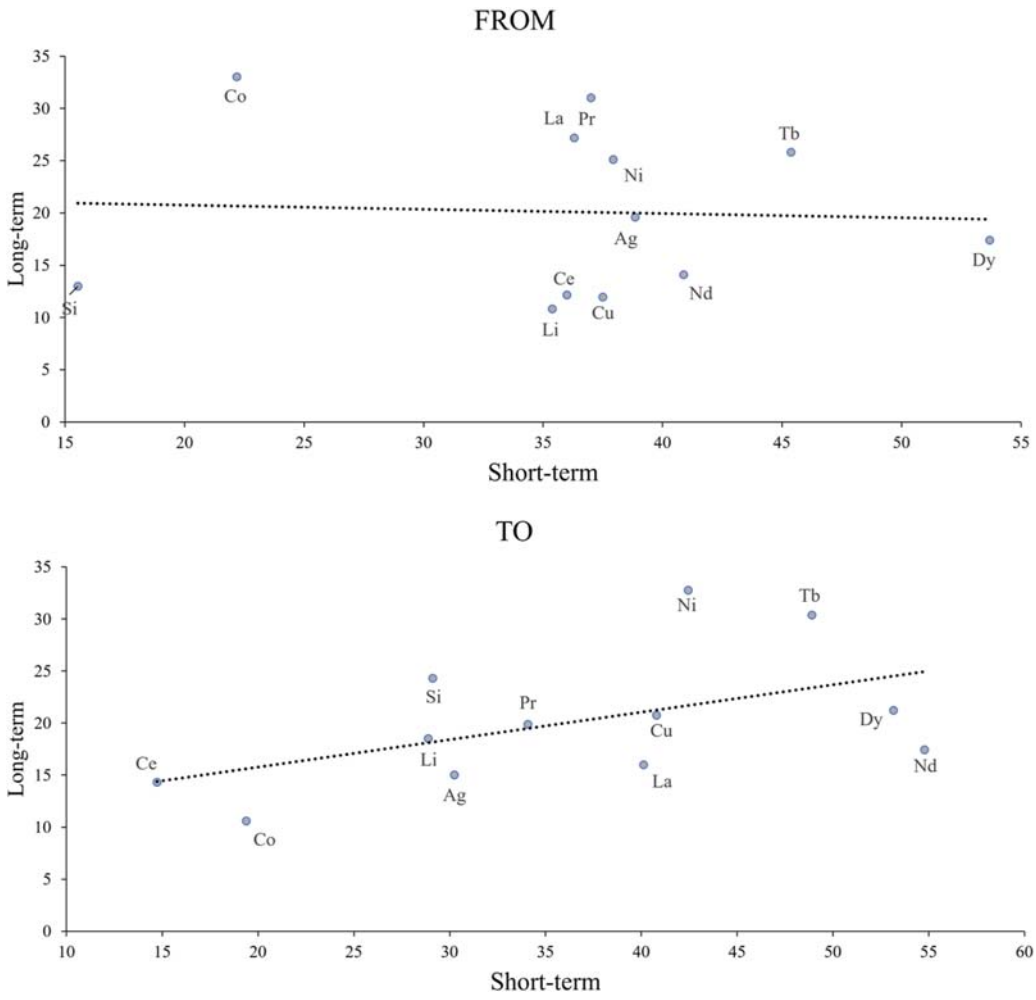


FIGURE 11
Return Total directional spillover in the short and long terms.

using the frequency-domain spillover index for analysis, since this index mainly analyses the spillover effect holistically from different frequency perspectives, this may result in a less sensitive response to the instantaneous fluctuations of neodymium than the time-domain spillover index. Specifically, skewness and kurtosis are statistical indicators used to measure the distribution pattern of data. Sharp short-term fluctuations often occur in the market. These fluctuations cause significant changes in skewness and kurtosis. If the frequency-domain spillover index focuses more on the long-term trend, it will "smooth out" these short-term and possibly exaggerated volatility effects, resulting in a final reflection that shows lower skewness and kurtosis spillover effects, i.e., short-term effects are masked by the long-term trend. That is, short-term effects are masked by long-term trends. This is the meaning of "suppressing short-term fluctuations". This makes the neodymium spillover effect in the short and long term relatively low, especially in the case of extreme market behaviours, the information becomes smoother, so the neodymium spillover effect under frequency-domain spillover index shows a weaker characteristic.

This phenomenon suggests that when conducting higher-order momentary risk spillover analysis, using the time-domain spillover index to monitor market volatility and risk transmission at a certain point in time, together with the frequency-domain spillover index to analyse its long-term trend and stability, can provide investors with a more forward-looking basis for investment decisions and identify potential risks and opportunities. Moreover, we compare the two types of spillover effects. If a market weakly influences others and receives minimal risk spillover, it becomes a suitable safe-haven investment. In this case, the stability of the market is more reliably verified.

With regard to the net spillover characteristics between new energy materials and rare earth elements (See Figures 15–18), the two markets can be discussed in four scenarios. First, combining long-term and short-term perspectives, similar to silicon for return risk and volatility risk, terbium for skewness risk, and cobalt for kurtosis risk, all of which have more significant long and short-term positive spillovers. Investors and policymakers should take full account of the risk shocks emanating from these markets or their strong persistence in risk management. Monitoring and analysis of market dynamics should be strengthened, investment portfolios should be flexibly adjusted, and a flexible risk management mechanism should be established to ensure that timely responses

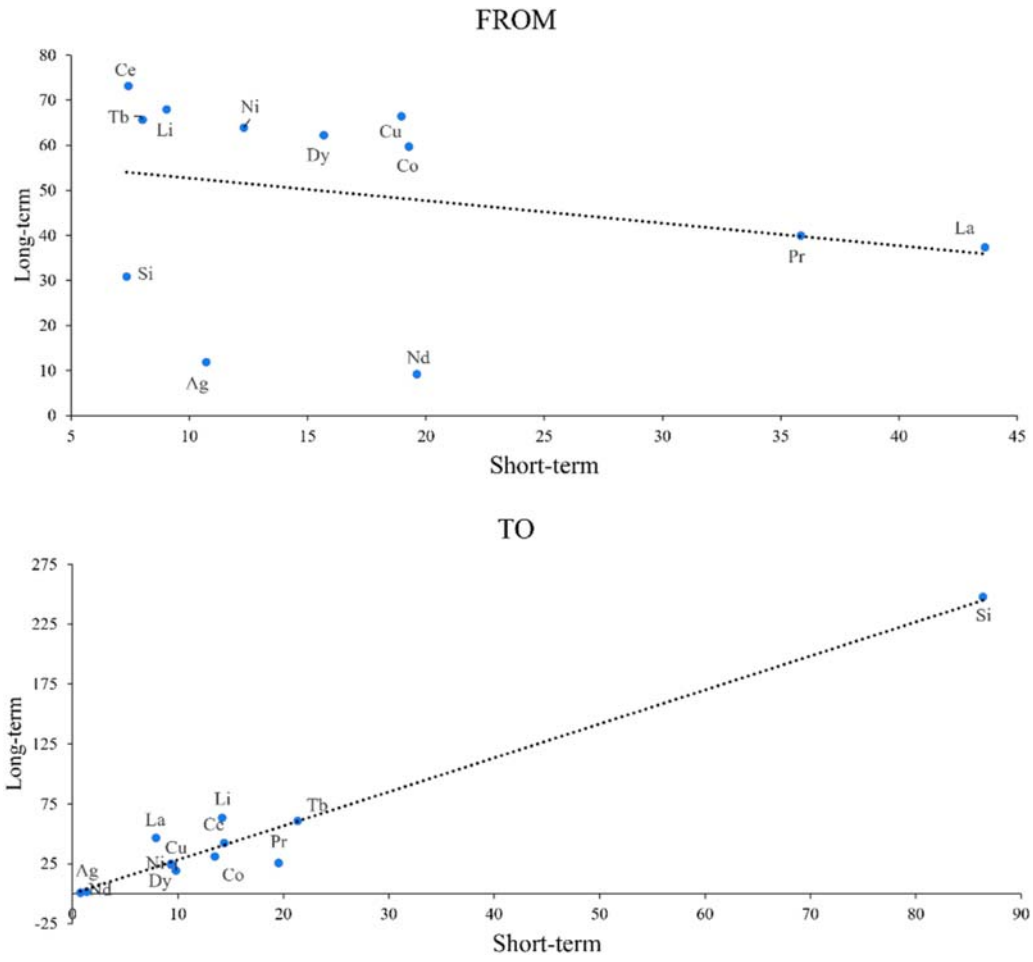


FIGURE 12
Volatility Total directional spillover in the short and long terms.

can be made in the event of both long-term and short-term risk spillovers.

With regard to the second category, these markets are net receivers in the short term and in the long term are in a more balanced position in terms of risk sending and receiving, demonstrating a distinctive market nature. For example, cerium for return risk, lanthanum for volatility and skewness risk, and silver for kurtosis risk. Investors should have the flexibility to respond to short-term market movements in such markets while focusing on long-term industry trends and fundamental changes in order to achieve effective risk control and optimise their investment portfolios when managing risk. The nature of such markets provides investors with the opportunity to trade on short-term volatility, while long-term stability is an important safeguard to protect the value of their investments.

Similarly, there are some markets that have significant risk-receiving characteristics over the long term, but are able to achieve a relative balance of risk in the short term. Examples include cobalt for return risk, nickel for volatility risk, and lithium for skewness risk. The ability of these markets to achieve a relative balance of risk in the short term is largely due to their efficient market mechanisms and the flexible response strategies of participants. For example, market participants used hedging instruments and dynamic asset allocation to cope with short-term volatility, thereby reducing potential risks. In addition, the transparency of relevant policy and market information facilitates quick adjustment of investment strategies. However, in the long term, these markets have again become notable risk takers due to their long-term supply and demand relationships and industry characteristics. The irreplaceability of metals such as lithium, cobalt and nickel in the new energy sector has led to the market's quest for long-term value to outweigh short-term volatility, and investors are willing to take on risk to capture potential long-term gains. Therefore, while short-term risks can be managed and regulated to maintain a balance, changes in market characteristics and demand structure in the long term make it a natural risk taker.

In addition to the markets mentioned above, we believe that if a market has significant negative net spillovers in both the short and long term, investors and regulators in such markets need to be particularly wary of risk transmission from price changes in other markets. Copper for volatility risk, neodymium for skewness risk, and nickel for kurtosis risk are typical. These metals markets are not only exposed to their own risk, but also receive a significant amount of risk impact from other markets, both in the short and

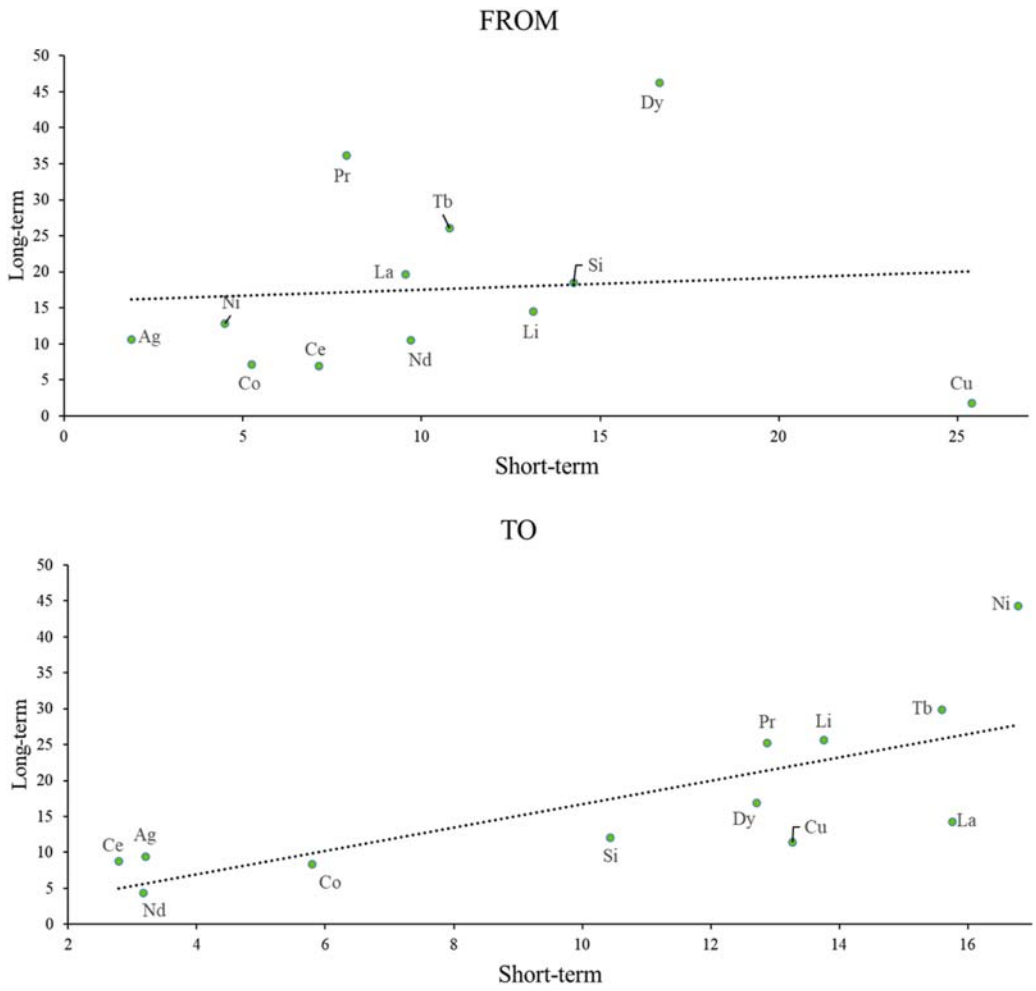


FIGURE 13
Skewness Total directional spillover in the short and long terms.

long term. Specifically, volatility-risky copper may absorb volatility risk factors from other markets as a result of economic uncertainty; skewness-risky neodymium shows sensitivity to asymmetric risk during market volatility, leading to the absorption of external risk; and kurtosis-risky nickel reflects the market's high sensitivity to the frequency of extreme events, which exacerbates the capture of risk from other markets.

6 Conclusion and discussion

By analysing the risk transfer between rare earth elements and new energy-related metal markets, this paper reveals the significant impact of co-production and co-consumption on market risk.

Based on the overall spillover indices of different order moments, the significant rise in the four types of risks mainly stems from the support of government industrial policies, resource control and increased global competition for resources. These policies raised market expectations of future returns while triggering increased short-term uncertainty, leading to frequent extreme market volatility and reflecting a complex risk environment.

Neodymium has been identified as a key major source of risk spillover in the extreme risk net spillover index and the maximum paired spillover index analyses. As the most consumed rare earth element, when neodymium price faces extreme risk, the risk factors may directly affect the neodymium market, which in turn affects the whole new energy industry chain. Therefore, it is recommended to prioritise the launch of neodymium metal futures contracts, giving full play to its advantages of high market maturity and stable industrial demand. In the specific implementation process, need to simultaneously establish supporting derivatives tools and standardised delivery system, through the cultivation of diversified market participants to enhance the liquidity, so as to build a perfect risk management mechanism. This initiative can effectively smooth out the impact of neodymium price fluctuations on the industrial chain, but also for the subsequent launch of other rare earth futures varieties to accumulate experience, and ultimately achieve the strategic goal of stabilising the rare earth market and ensuring the healthy development of the industry.

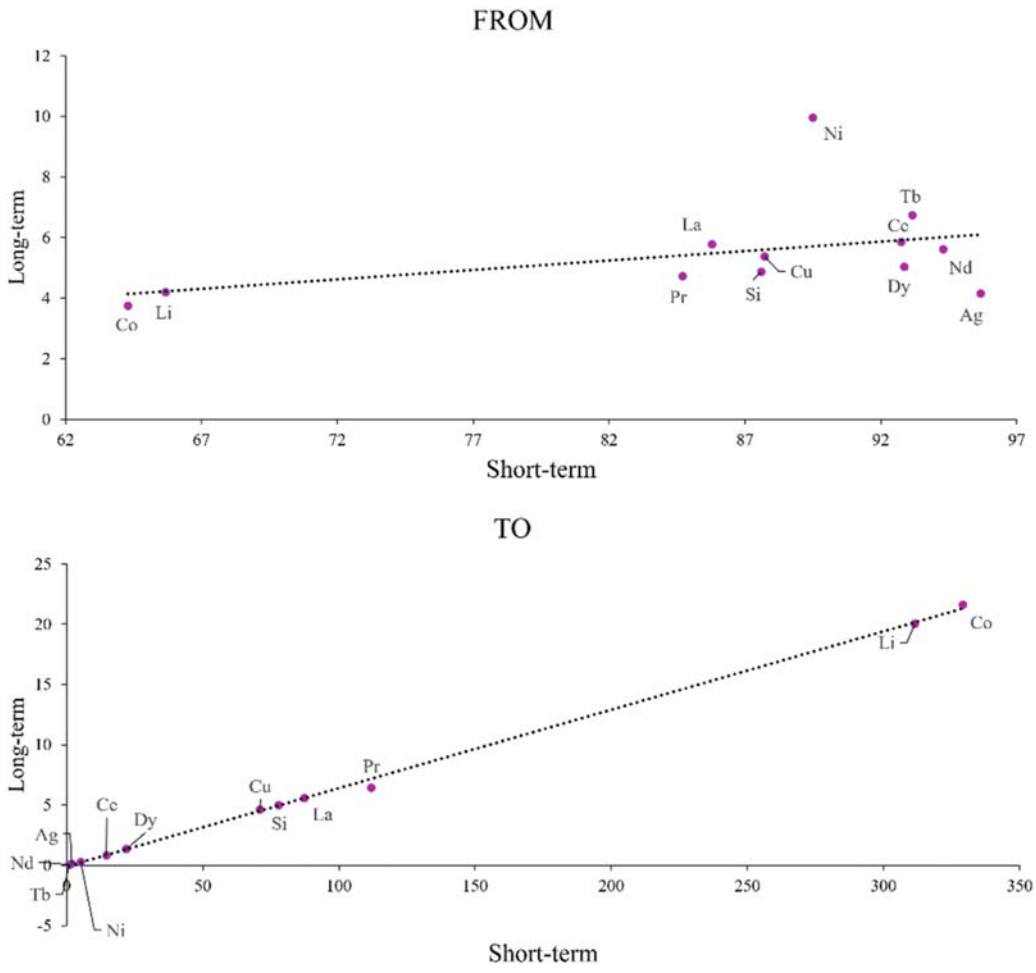


FIGURE 14
Kurtosis Total directional spillover in the short and long terms.

The Maximum Pairwise Spillover Index reveals the interactions and risk transfer in the market for new energy-related metals and rare earth elements. In the skewness and kurtosis risk premium analyses, the maximum risk spillover value is significantly higher than return and volatility, suggesting a strong asymmetric price volatility risk-transfer effect in the market as well as anomalous volatility connectivity in the face of external shocks. All of the above strongly correlated metal combinations have co-production and co-consumption relationships. This suggests that the two types of production-consumption relationships are highly susceptible to strong potential interactions in the corresponding metal markets and that policymakers should pay close attention to metals that have such relationships as a result of technological developments.

In the analysis of the evolution of spillover networks, the details of the evolution of volatility spillover networks are the most interesting to explore compared to other risk spillover networks that are fixed in form. In particular, rare earth markets tend to evolve as risk emitters, while new energy-related metal markets are more likely to evolve as risk receivers. This situation highlights the key position of rare earths in the new energy industry chain, coupled with the fact that the rare earth element market is more vulnerable to policy regulation and international market competition, its market instability and risk transmission has the potential risk of triggering a chain reaction, which requires the industry chain participants to pay close attention to market dynamics, strengthen risk management and cooperation, in order to ensure the stability and sustainable development of the new energy industry chain.

For downstream industries such as wind power and electric vehicles, this study suggests that enterprises focus on the price fluctuation risk of metal markets that are strongly correlated with rare earth elements. For example, the wind power industry relies on NdFeB permanent magnet materials, while the electric vehicle industry chain is highly sensitive to the demand for cobalt and copper, so companies can hedge the risk of raw material price fluctuations by signing long-term supply agreements or using futures tools.

In the comparative analysis of long- and short-term market linkages for each order of moment risk, the relationship between long- and short-term spillovers for volatility risk is more pronounced, with acceptance spillovers showing a negative

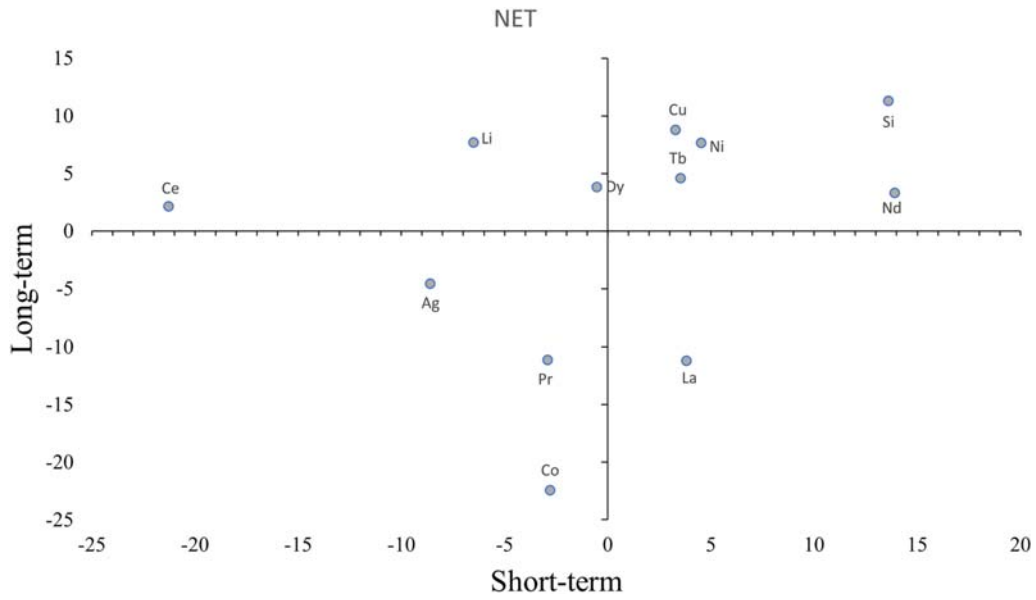


FIGURE 15
Return NET spillover in the short and long terms.

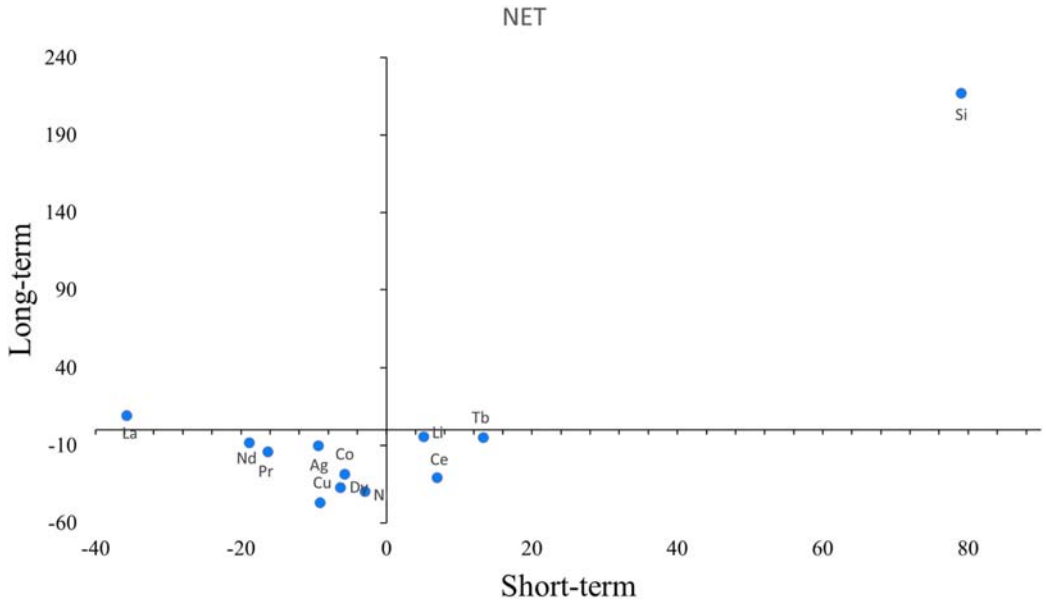


FIGURE 16
Volatility NET spillover in the short and long terms.

correlation and external spillovers showing a significant positive correlation. For both skewness and kurtosis risk, there is a clear positive long- and short-term correlation for the external spillover, and when the market has a strong short-term external impact, the long-term impact is also strong. In terms of net spillover characteristics, for markets with significant positive spillovers in both the long and short term, investors and policymakers need to focus on risk shocks from these markets. For markets that are net receivers in the short run and where the net risk spillover is close to zero in the long run. Investors can adopt active investment strategies, such as capturing short-term upside and liquidity opportunities, while flexibly adjusting their positions in response to market volatility. The other type of market is one that can achieve a relatively balanced risk profile in the short term, but become a significant risk taker in the long term. Investors should implement effective

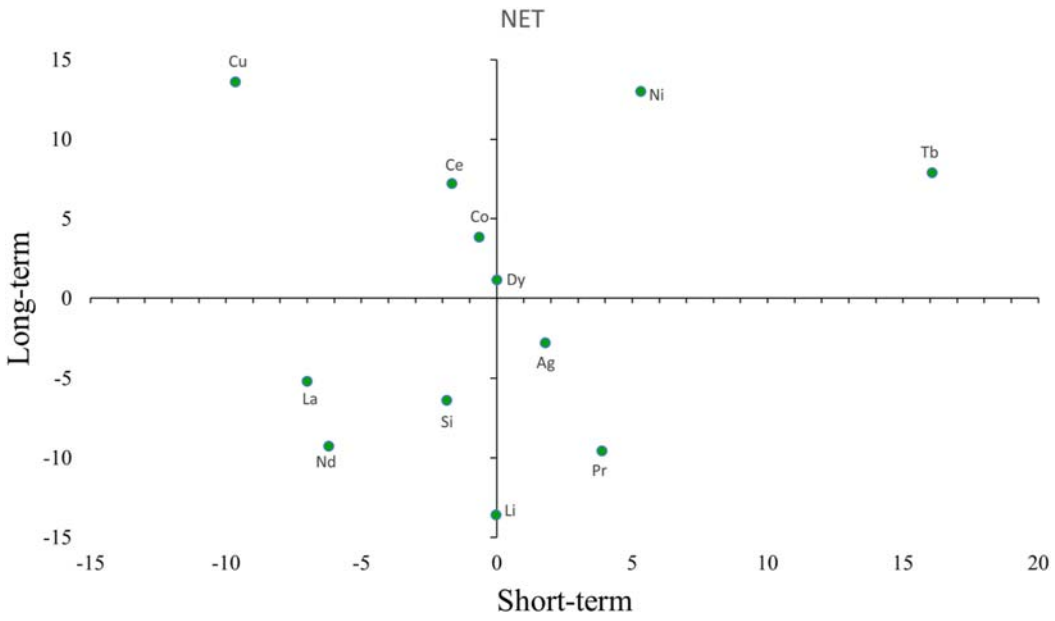


FIGURE 17
Skewness NET spillover in the short and long terms.

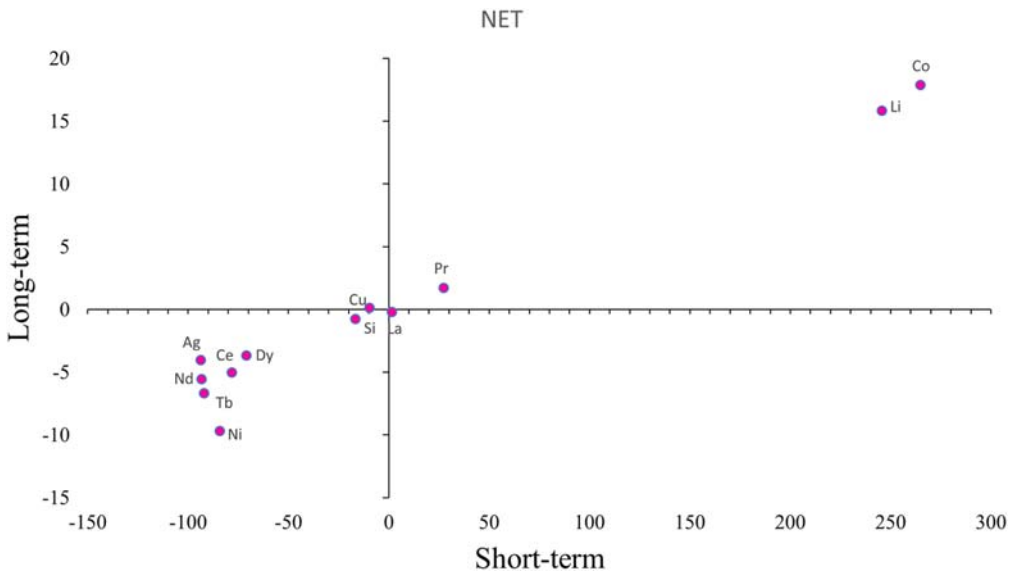


FIGURE 18
Kurtosis NET spillover in the short and long terms.

risk management measures, including asset diversification, hedging strategies and dynamic portfolio adjustments, to reduce the risk of long-term volatility and ensure the sustainability and stability of their investments. Finally, for markets with significant negative net spillovers in both the short and long term, investors and regulators need to be particularly vigilant against risk transmission from price changes in other markets. These findings provide investors with an important basis for risk management, helping to optimize portfolios and protect investment value.

In order to obtain more in-depth conclusions, future research could consider constructing quantitative indicators of policy impacts and geopolitics. In this way, the shock transmission paths of both can be comprehensively analysed.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Requests to access these datasets should be directed to Hao Nong, 231601011@csu.edu.cn.

Author contributions

RZ: Visualization, Data curation, Methodology, Formal Analysis, Writing – review and editing, Software, Supervision, Writing – original draft, Resources, Conceptualization, Validation. HN: Conceptualization, Writing – review and editing, Formal Analysis, Writing – original draft, Data curation, Methodology. QC: Resources, Validation, Writing – review and editing, Conceptualization, Formal Analysis. LS: Conceptualization, Writing – original draft, Funding acquisition, Validation, Project administration, Formal Analysis. HZ: Software, Conceptualization, Resources, Formal Analysis, Writing – review and editing. YZ: Writing – review and editing, Software, Conceptualization, Validation, Formal Analysis.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors wish to express their gratitude for the support provided by National Science and Technology Major Project (2024ZD1002002), the National Natural Science Foundation of China (No. 72204273).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1616099/full#supplementary-material

References

1. Vakulchuk R, Overland I. Central Asia is a missing link in analyses of critical materials for the global clean energy transition. *One Earth* (2021) 4:1678–92. doi:10.1016/j.oneear.2021.11.012

2. Apergis E, Apergis N. The role of rare earth prices in renewable energy consumption: the actual driver for a renewable energy world. *Energy Econ* (2017) 62:33–42. doi:10.1016/j.eneco.2016.12.015

3. Madaleno M, Taskin D, Dogan E, Tzeremes P. A dynamic connectedness analysis between rare earth prices and renewable energy. *Resour Pol* (2023) 85:103888. doi:10.1016/j.resourpol.2023.103888

4. Xu JL, Li JH, Charles V, Zhao X. Evolution of the rare earth trade network: a perspective of dependency and competition. *Geosci Front* (2024) 15:101653. doi:10.1016/j.gsf.2023.101653

5. Gao W, Wei J, Zhang H, Zhang H. The higher-order moments connectedness between rare earth and clean energy markets and the role of geopolitical risk: new insights from a TVP-VAR framework. *Energy* (2024) 305:132280. doi:10.1016/j.energy.2024.132280

6. Afflerbach P, Fridgen G, Kellerº R, Rathgeber AW, Strobel F. The by-product effect on metal markets – new insights to the price behavior of minor metals. *Resour Pol* (2014) 42:35–44. doi:10.1016/j.resourpol.2014.08.003

7. Zhang H, Li Z, Song H, Gao W. Insight into clean energy market's role in the connectedness between joint-consumption metals. *Energy* (2024) 302:131831. doi:10.1016/j.energy.2024.131831

8. Shammugam S, Rathgeber A, Schlegl T. Causality between metal prices: is joint consumption a more important determinant than joint production of main and by-product metals? *Resour Pol* (2019) 61:49–66. doi:10.1016/j.resourpol.2019.01.010

9. Sun X. Supply chain risks of critical metals: sources, propagation, and responses. *Front Energy Res* (2022) 10:957884. doi:10.3389/fenrg.2022.957884

10. Nong H, Huang Q. Total, quantile, and frequency risk transmission among metal commodities. *Int J Finance and Econ* (2024). doi:10.1002/ijfe.3017

11. Su H, Wu Q, Zhou N, Zheng Y. Dynamic spillover of upstream primary metals and by-product metals market and its impact from the downstream new energy vehicles market. *Mineral Econ* (2024) 38:77–95. doi:10.1007/s13563-024-00456-9

12. Pata UK, Erdogan S, Pata SK. Response of renewable energy minerals to energy policy and technological advancement: an assessment of the Russian energy crisis. *J Environ Manage* (2024) 370:122608. doi:10.1016/j.jenvman.2024.122608

13. Liu FY, Su CW, Qin M, Lobont OR. Winner or loser? The bidirectional impact between geopolitical risk and energy transition from the renewable energy perspective. *Energy* (2023) 283:129174. doi:10.1016/j.energy.2023.129174

14. Nwonye NG, Onuselogu OCO, Anisiuba CA, Ezeaku HC, Egbo OP. Dynamics of green metal price volatility in times of geopolitical tensions: effects of oil price shocks and carbon emissions futures. *J Clean Prod* (2023) 412:137383. doi:10.1016/j.jclepro.2023.137383

15. Dou SQ, Zhu YG, Liu JY, Xu DY. The power of mineral: shock of the global supply chain from resource nationalism. *World Dev* (2024) 184:106758. doi:10.1016/j.worlddev.2024.106758

16. Baldi L, Peri M, Vandone D. Clean energy industries and rare earth materials: economic and financial issues. *Energy Policy* (2014) 66:53–61. doi:10.1016/j.enpol.2013.10.067

17. Reboredo JC, Ugolini A. Price spillovers between rare earth stocks and financial markets. *Resour Pol* (2020) 66:101647. doi:10.1016/j.resourpol.2020.101647

18. Gao Y, Liu X. Time and frequency spillovers and drivers between rare earth and energy, metals, green, and agricultural markets. *North Am J Econ Finance* (2024) 72:102128. doi:10.1016/j.najef.2024.102128

19. Ding Q, Huang J, Chen J. Time-frequency spillovers and the determinants among fossil energy, clean energy and metal markets. *Energy J* (2023) 44:259–86. doi:10.5547/01956574.44.2.qdin

20. Kamal E, Bouri E. Dependence structure among rare earth and financial markets: a multiscale-vine copula approach. *Resour Pol* (2023) 83:103626. doi:10.1016/j.resourpol.2023.103626

21. Zheng B, Zhang YW, Qu F, Geng Y, Yu H. Do rare earths drive volatility spillover in crude oil, renewable energy, and high-technology markets? —a wavelet-based BEKK-GARCH-X approach. *Energy* (2022) 251:123951. doi:10.1016/j.energy.2022.123951

22. Weng Z, Haque N, Mudd GM, Jowitt SM. Assessing the energy requirements and global warming potential of the production of rare earth elements. *J Clean Prod* (2016) 139:1282–97. doi:10.1016/j.jclepro.2016.08.132

23. Balaram V. Rare earth elements: a review of applications, occurrence, exploration, analysis, recycling, and environmental impact. *Geosci Front* (2019) 10:1285–303. doi:10.1016/j.gsf.2018.12.005

24. Dushyantha N, Batapola N, Ilankoon I, Rohitha S, Premasiri R, Abeysinghe B, et al. The story of rare earth elements (REEs): occurrences, global distribution, genesis, geology, mineralogy and global production. *Ore Geology Rev* (2020) 122:103521. doi:10.1016/j.oregeorev.2020.103521

25. Cuadros-Muñoz J-R, Jimber-Del-Río J-A, Sorhegui-Ortega R, Zea-De La Torre M, Vergara-Romero A. Contribution of rare earth elements is key to the economy of the future. *Land* (2024) 13:1220. doi:10.3390/land13081220

26. Binnemans K, Jones PT, Müller T, Yurramendi L. Rare earths and the balance problem: how to deal with changing markets? *J Sustainable Metall* (2018) 4:126–46. doi:10.1007/s40831-018-0162-8

27. Nakagawa K, Uchiyama Y. GO-GJRSK model with application to higher order risk-based portfolio. *Mathematics* (2020) 8:1990. doi:10.3390/math8111990

28. Antonakakis N, Chatziantoniou I, Gabauer D. Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *J Risk Financial Manage* (2020) 13:84. doi:10.3390/jrfm13040084

29. Diebold FX, Yilmaz K. Better to give than to receive: predictive directional measurement of volatility spillovers. *Int J Forecast* (2012) 28:57–66. doi:10.1016/j.ijforecast.2011.02.006

30. Ding Q, Huang JB, Chen JY. Dynamic and frequency-domain risk spillovers among oil, gold, and foreign exchange markets: evidence from implied volatility. *Energy Econ* (2021) 102:105514. doi:10.1016/j.eneco.2021.105514

31. Baruník J, Křehlík T. Measuring the frequency dynamics of financial connectedness and systemic risk. *J Financial Econom* (2018) 16:271–96. doi:10.1093/jjfinec/nby001

32. Blanco MI. The economics of wind energy. *Renew Sustain Energ Rev* (2009) 13:1372–82. doi:10.1016/j.rser.2008.09.004

33. Buchholz P, Brandenburg T. Demand, supply, and price trends for mineral raw materials relevant to the renewable energy transition wind energy, solar photovoltaic energy, and energy storage. *Chem Ingenieur Technik* (2018) 90:141–53. doi:10.1002/cite.201700098

34. Månberger A, Stenqvist B. Global metal flows in the renewable energy transition: exploring the effects of substitutes, technological mix and development. *Energy Policy* (2018) 119:226–41. doi:10.1016/j.enpol.2018.04.0568

35. Alves Dias P, Bobba S, Carrara S, Plazzotta B. *The role of rare earth elements in wind energy and electric mobility*. Luxembourg: European Commission (2020).

36. Jones B, Elliott RJ, Nguyen-Tien V. The EV revolution: the road ahead for critical raw materials demand. *Appl Energy* (2020) 280:115072. doi:10.1016/j.apenergy.2020.115072

37. Ortego A, Calvo G, Valero A, Iglesias-Embil M, Valero A, Villacampa M. Assessment of strategic raw materials in the automobile sector. *Resour Conservation Recycling* (2020) 161:104968. doi:10.1016/j.resconrec.2020.104968

38. Calvo G, Valero A. Strategic mineral resources: availability and future estimations for the renewable energy sector. *Environ Dev* (2022) 41:100640. doi:10.1016/j.envdev.2021.100640

39. Farina A, Anctil A. Material consumption and environmental impact of wind turbines in the USA and globally. *Resour Conservation Recycling* (2022) 176:105938. doi:10.1016/j.resconrec.2021.105938

40. Gielen D, Lyons M. Critical materials for the energy transition: rare earth elements. *Int Renew Energy Agency* (2022) 1–48.

41. Peiró LT, Martin N, Méndez GV, Madrid-López C. Integration of raw materials indicators of energy technologies into energy system models. *Appl Energy* (2022) 307:118150. doi:10.1016/j.apenergy.2021.118150

42. Maani T, Mathur N, Rong CB, Sutherland JW. Estimating potentially recoverable Nd from end-of-life (EoL) products to meet future US demands. *Resour Conservation Recycling* (2023) 190:106864. doi:10.1016/j.resconrec.2023.106864

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 18 June 2025
DOI 10.3389/fphy.2025.1599993

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Tingwei Fang,
Capital University of Economics and Business, China
Tao Wu,
Chengdu University of Technology, China

*CORRESPONDENCE
Liang Wu,
100407366@gznu.edu.cn

RECEIVED 25 March 2025
ACCEPTED 30 May 2025
PUBLISHED 18 June 2025

CITATION
Lei L, Wang Y, Wu L and Mao Y (2025) Spillover effects in stock index returns within Chinese oil industry chain.
Front. Phys. 13:1599993.
doi: 10.3389/fphy.2025.1599993

COPYRIGHT
© 2025 Lei, Wang, Wu and Mao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Spillover effects in stock index returns within Chinese oil industry chain

Likun Lei[1], Yu Wang[1], Liang Wu[2]* and Yufeng Mao[1]

[1]School of Application and Economics, Guizhou University of Finance and Economics, Guiyang, China, [2]School of Economics and Management, Guizhou Normal University, Guiyang, Guizhou, China

With the global energy market becoming increasingly tight and the rapid development of Chinese oil industry, the transmission of information and mutual influence among sectors have become increasingly prominent. In light of this, this paper focuses on the three major sectors within Chinese oil industry chain from 2014 to 2024: oil extraction, oil processing, and oil trading. By employing the TVP-VAR-DY and BK models, we analyze the return spillover effects and dynamic correlations among the upstream, midstream, and downstream sectors of the oil industry chain. The results indicate that there are significant return spillover effects among these three sectors, with their price fluctuations not existing independently but rather influencing and being influenced by each other. Specifically, the oil trading sector predominantly acts as a spillover recipient, while the oil processing sector consistently serves as a net spillover transmitter across the three markets. Furthermore, as the frequency of spillovers increases, the spillover index among the three markets gradually decreases. Notably, when confronted with extreme event shocks, the overall total spillover level of the oil industry system significantly amplifies.

KEYWORDS

oil industry chain, spillover effect, TVP-VAR-DY model, BK model JEL classification: C22, C53, G11, G17

1 Introduction

Oil, hailed as the "lifeblood of modern industry," is a vital lifeline for a country's economic development and industrial production, with its market dynamics and the interactions among various segments of the industrial chain increasingly becoming the focus of attention in academia and industry (see, e.g., [1–6]). Within this chain, industries such as oil extraction, processing, and trade are intricately interconnected and mutually influential. Oil extraction, as the upstream segment, directly impacts the production costs and market supply of downstream oil processing enterprises through its output and costs. Oil processing, in turn, produces a variety of petrochemical products by refining crude oil to meet market demands. Meanwhile, oil trade, bridging domestic and international markets, is an indispensable link in the oil industry chain.

With the changes in the global economic situation and the adjustment of energy structure, the spillover effects within the oil industry chain have become increasingly pronounced. Since the outbreak of the 2008 financial crisis, the crude oil market has been in a persistent state of vulnerability (see, e.g., [7]). Notably, in recent years, the convergence of major unexpected events like the 2019 pandemic and the Russia-Ukraine conflict has significantly reduced global oil demand, undoubtedly exacerbating the instability and

unpredictability of crude oil market prices (see, e.g., [8]). In this context, the spillover effects and dynamic correlations between various industries within the oil industry chain are more significant. On the one hand, price fluctuations in domestic and international oil markets exert direct or indirect impacts on all segments of the oil industry chain. On the other hand, internal price fluctuations within the oil industry chain can also be transmitted to other financial markets through trade, investment, and other channels, triggering broader market volatility. This spillover effect not only escalates risks within the oil industry chain but also poses greater challenges to related enterprises and investors. Therefore, analyzing the spillover effects within Chinese oil industry chain not only uncovers the interactive mechanisms of price fluctuations among various sectors within the industry chain but also provides policymakers with a scientific basis for risk management and market regulation. Furthermore, it aids investors in optimizing their investment portfolios, reducing market risks in the face of extreme events, and fostering the stable development of the oil industry chain.

The aims of this study are to comprehensively analyzes the price spillover effects and their time-varying characteristics among the oil extraction, processing, and trade sectors within Chinese oil industry chain. To achieve these aims, the paper primarily addresses the following two questions: Firstly, how to accurately depict and quantify the return spillover relationships and net spillover effects among these three sectors? Secondly, how to uncover the transmission mechanisms and dynamic changes of these spillover effects across different frequency domains (short-term, medium-term, and long-term)? To address these questions, this study employs stock index data from the oil extraction, oil processing, and oil trade sectors within the Shenwan Tertiary Industry Index. By employing a combined method of the TVP-VAR-DY model and the BK model, we aim to capture and deeply uncover the dynamic variation characteristics of the return spillover effects among different sectors within the petroleum industry chain, as well as the evolution of these effects over time.

The contributions of this article lie in the following three aspects: firstly, it integrates the oil extraction, oil processing, and oil trade industries in Chinese oil industry chain into a unified research framework, and comprehensively analyzes the return spillover relationship and net spillover relationship among the three through the TVP-VAR-DY model. Secondly, the BK model was introduced to measure the dynamic spillover effects among the three major industries in the oil industry chain from a combined perspective of time and frequency domains. We analyze the transmission mechanisms of price signals across different frequency domains in depth, revealing the mutual influence of price fluctuations in the three industries in the short, medium, and long term. Finally, we analyze in detail the net directional and bidirectional spillover effects of oil extraction, oil processing, and oil trade industries in different frequency domains, and drew spillover network diagrams to visually display the different roles of the three industries in spillover propagation, and tracked the dynamic changes of these roles.

The remainder of this paper is structured as follows: Section 2 presents a literature review, summarizing previous research on spillover effects within crude oil markets. Section 3 introduces the research methodology, including the application of the TVP-VAR-DY model and the BK model. Section 4 conducts an empirical analysis, presenting the detailed research findings. Finally, Section 5 concludes the paper, offering recommendations and future research directions.

2 Literature review

Current scholarship has extensively investigated spillover effects in crude oil markets. Researchers have utilized various econometric approaches to examine inter-market relationships within international crude oil markets. According to the market linkage theory, prices across different markets exhibit characteristics of mutual influence and coordinated fluctuations due to factors such as economic interconnections and trade activities. Chen et al. [9], Elder et al. [10], Liao et al. [11], Caporin et al. [12] empirically reveal that there is a significant correlation and co-integration between WTI and Brent crude oil market, and Hammoudeh et al. [13] further demonstrate significant spillover effects between crude oil and refined product markets, including various petroleum products such as heating oil and gasoline, Additionally, Ji and Fan [14] find a lead-lag relationship among major international crude oil markets (WTI, Brent, Dubai, Tapis, and Nigeria), providing a new perspective for understanding the dynamics of the global crude oil market. Leong [15] focuses on the causal relationship between BRENT crude oil futures returns and crude oil related industry stock indices, as well as the dynamic correlation and risk spillover effects between international oil prices and domestic oil related industry stock markets, further confirming the heterogeneous impact of oil price fluctuations on industry markets.

The interaction between futures and spot markets remains a central topic in energy finance research. Studies by Huang et al. [16], Lei and Yong [17], Lee and Zeng [18], Wang and Wu [19], Mehrara and Hamldar [20], Chen et al. [21], Chang and Lee [22], Klein [23] substantially confirm the substantial correlation and spillover effects between futures and spot markets for WTI and Brent crude oil. Furthermore, Magkonis and Tsouknidis [24] extend this understanding by revealing similar spillover effects between futures and spot prices of crude oil commodities across various markets, thereby deepening our comprehension of crude oil market integration. With the advancement of research, scholars have increasingly focused on the price transmission mechanisms within the petroleum industry chain. Complex Systems Theory and Sociophysics posit that financial markets constitute dynamic networks formed through nonlinear interactions among a multitude of heterogeneous agents (e.g., investors, institutions, and information nodes), emphasizing that market fluctuations stem from the superposition of multiple micro-level behavioral feedback loops rather than being driven by singular factors. The petroleum industry chain represents a quintessential complex network system, where price fluctuations are not transmitted through simple linear relationships but are instead influenced by multiple factors, such as market structure, policy interventions, and investor behaviors. For instance, Guo et al. [25] select 10 oil products from three crucial links—oil production, consumption, and market transactions—within the oil industry chains across global markets. Employing network motif analysis, they explored the price lead-lag relationships among these products. Their findings reveal the presence of price lead-lag relationships both horizontally across regions and vertically

within the industry chain. Qi et al. [26] comprehensively analyzed the price spillover relationships between major commodities within the Chinese petroleum industry chain by constructing a multidimensional spillover network, revealing the transmission path and mediating role of crude oil price fluctuations within the industry chain.

With the rapid development of Chinese economy and the continuous growth of energy consumption, the Chinese crude oil market has become increasingly important in the international crude oil market. Scholars have conducted extensive research on the correlation between the Chinese crude oil market and the international crude oil market. These studies all show that there is a significant correlation and spillover effect between Chinese crude oil prices and international crude oil prices (see, e.g., [3, 27–31]).

Compared with existing studies, this paper contributes to the literature in three key aspects. Firstly, while most existing research concentrates on price dynamics in the overall market or specific markets, this study provides a more targeted and in-depth analysis by focusing on price transmission mechanisms across different segments of Chinese petroleum industry chain. Second, methodologically, we introduce an enhanced time-varying parameter vector autoregression (TVP-VAR) model, which allows for more precise identification of nonlinear transmission effects. Third, in terms of theoretical contribution, our findings not only confirm the vertical transmission of price information along the industry chain but also reveal the critical mediating role of petroleum processing—a novel insight that enriches the understanding of price signal propagation in capital markets. By quantifying the micro-level price transmission mechanisms within Chinese petroleum industry, this study offers policymakers and corporate decision-makers more actionable insights, underscoring the novelty and significance of our research.

3 Methodology

3.1 TVP-VAR-DY spillover index model

To further explore the spillover effects and time-varying characteristics of the three key links in the oil industry chain, namely, extraction, processing, and trade, this article employs the TVP-VAR-DY model proposed by Antonakakis et al. [32], which combines the time-varying parameter vector autoregressive (TVP-VAR) model with the spillover index method (DY) based on generalized variance decomposition proposed by Diebold and Yilmaz [33], Diebold and Yılmaz [34], It provides an effective tool for measuring the relationship and influences between financial markets (see, e.g., [35–40]). In the specific research process, the TVP-VAR model is constructed by Equations 1, 2:

$$Y_t = \beta_t Y_{t-1} + \varepsilon_t, \varepsilon_t \sim N\left(0, \sum_t\right) \tag{1}$$

$$vec(\beta_t) = vec(\beta_{t-1}) + v_t, v_t \sim N(0, R_t) \tag{2}$$

where Y_t, Y_{t-1}, and ε_t are all $N\times1$ dimensional vectors, which represent the price series and corresponding errors of the stock indices of the oil exploitation, oil processing and oil trading industries in the t and t-1periods, respectively. β_t and $\sum_t$ are $N\times N$ dimensional vector matrices representing the time-varying VAR coefficient and the time-varying variance-covariance, respectively. Additionally, $vec(\beta_t)$ and v_t are the $N^2\times1$ dimensional vector matrices.

Furthermore, Koop et al. [41], Pesaran and Shin [42] introduced the concept of Generalized Forecast Error Variance Decomposition (GFEVD) based on the Wold theorem. To effectively apply this concept, it is necessary to transform the TVP-VAR model into a TVP-VMA process (Equation 3):

$$Y_t = \sum_{i=1}^{p} \beta_{it} Y_{t-i} + \varepsilon_t = \sum_{j=1}^{\infty} B_{jt} \varepsilon_{j-t} \tag{3}$$

where B_{jt} is a $N\times N$ dimensional vector matrix. Next, GFEVD is calculated, which is used to quantify the variance share of a variable in predicting the error of another variable. It reflects the degree to which a variable's change is influenced by itself or other variables in the system, and is calculated as Equations 4, 5:

$$\theta_{ij,t}^{H} = \frac{\sum_{jj,t}^{-1} \sum_{h=0}^{H-1} \left(e_{i,t}' B_{h,t} \sum_t e_j \right)^2}{\sum_{h=0}^{H-1} \left(e_{i,t}' B_{h,t} \sum_t B_{h,t}' e_{i,t} \right)} \tag{4}$$

$$\tilde{\theta}_{ij,t}^{H} = \frac{\theta_{ij,t}^{H}}{\sum_{j=1}^{N} \theta_{ij,t}^{H}} \tag{5}$$

where H represents the prediction period, e_i and e_j are the selection vectors (the ith element in e_i is 1, and the remaining elements are 0; the same applies to e_j). $\tilde{\theta}_{ij,t}^{H}$ represents the impact of variable j on variable i in the H -step prediction variance. Since the $\theta_{ij,t}^{H}$ is non orthogonalized, the sum of the contributions of the prediction error method may not necessarily be equal to 1, that is, the sum of all rows in $\theta_{ij,t}^{H}$ may not necessarily be 1. Therefore, standardization of $\theta_{ij,t}^{H}$ will be carried out to ensure that $\Sigma_{j=1}^{N}\theta_{ij,t}^{H} = 1$, and $\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}\theta_{ij,t}^{H} = N$. Chatziantoniou and Gabauer [43] confirmed through Monte Carlo simulation that the range of values for the traditional total spillover index is [0, (m-1)/m], and therefore adjusted it to the following expression (Equation 6):

$$S_t^H = \frac{\sum_{i=1,j=1}^{N} \tilde{\theta}_{ij,t}^{H}}{\sum_{i=1,j=1}^{N} \tilde{\theta}_{ij,t}^{H}} \times 100 = \frac{i \neq j}{N} \times 100 \tag{6}$$

Based on the prediction error variance matrix constructed above, the following spillover indices can be obtained:

The spillover from market i to other markets j is defined as Equation 7:

$$TO_{it}(H) = \Sigma_{i=1,i\neq j}^{N} \tilde{\theta}_{ij,t}^{H} \tag{7}$$

Market i receiving spillover from other markets j is given by Equation 8:

$$FROM_{it}(H) = \Sigma_{j=1,j\neq i}^{N} \tilde{\theta}_{ij,t}^{H} \tag{8}$$

The net spillover effect of market i on other markets j can be calculated as Equation 9:

$$NET_{it}(H) = TO_{it}(H) - FROM_{it}(H) \tag{9}$$

The net pairing spillover index $NPDC$ is given by Equation 10:

$$NPDC_{ij,t} = \tilde{\theta}^{H}_{ij,t} - \tilde{\theta}^{H}_{ji,t} \tag{10}$$

3.2 Frequency-domain spillover index model

When a market faces price shocks from other markets, its response and price spillover effects may exhibit different durations, that is, different frequency performance. Baruník and Křehlík [44] argue that the DY spillover index method is apply at the overall level, which may cover spillovers of different frequencies. Therefore, researchers adopt the spectral representation method of covariance. This method obtains the frequency response function through the Fourier transform of coefficients, which can then be used to obtain the generalized variance decomposition in the frequency domain. Specifically, at a given frequency ω, the calculation process of generalized prediction error variance decomposition is as follows:

According to the spectral decomposition method of Stiassny [45], the frequency response function $\psi(e^{-iw}) = \Sigma_{h=0}^{\infty} e^{-iwh}\psi_h$ is first transform into its TVP-VMA (∞) form using Fourier transform (Equation 11):

$$S_x(\omega) = \Sigma_{h=-\infty}^{\infty} E\left(x_t x'_{t-h}\right) e^{-iwh} = \psi_t\left(e^{-iwh}\right)\Sigma\Psi'_t\left(e^{iwh}\right) \tag{11}$$

Frequency GFEVD is a combination of spectral density and GFEVD, which also requires normalization in the time domain as shown in Equations 12, 13:

$$\theta_{ij,t}(\omega) = \frac{\left(\Sigma_t\right)_{jj}^{-1}\left|\Sigma_{h=0}^{\infty}\left(\Psi_t\left(e^{-iwh}\right)\Sigma_t\right)_{ij,t}\right|^2}{\Sigma_{h=0}^{\infty}\left(\psi_t\left(e^{-iwh}\right)\Sigma_t\Psi_t\left(e^{iwh}\right)\right)_{ii}} \tag{12}$$

$$\tilde{\theta}_{ij,t}(\omega) = \frac{\theta_{ij,t}(\omega)}{\Sigma_{j=1}^{N}\theta_{ij,t}(\omega)} \tag{13}$$

Where $\tilde{\theta}^{H}_{ij,t}(\omega)$ represents the impact of the ith variable on the jth variable at frequency ω. This single frequency can be used to measure short-term, medium-term, and long-term spillover effects. For a given frequency band, where: $d = (a, b), a, b \in (-\pi, \pi), a < b$. The frequency band d can be define as Equation 14:

$$\tilde{\theta}_{ij,t}(d) = \int_a^b \tilde{\theta}_{ij,t}(\omega)d\omega \tag{14}$$

The total revenue spillover within frequency band d can be calculate using the following equations (from Equations 15–19):

$$NPDC_{ij,t}(d) = \tilde{\theta}_{ij,t}(d) - \tilde{\theta}_{ji,t}(d) \tag{15}$$

$$TO_{it}(d) = \Sigma_{i=1,i\neq j}^{N}\tilde{\theta}_{ij,t}(d) \tag{16}$$

$$FROM_{it}(d) = \Sigma_{j=1,j\neq i}^{N}\tilde{\theta}_{ij,t}(d) \tag{17}$$

$$NET_{it}(d) = TO_{it}(d) - FROM_{it}(d) \tag{18}$$

$$TCI_t(d) = \frac{N}{N-1}\Sigma_{i=1}^{N}TO_{it}(d) = \frac{N}{N-1}\Sigma_{i=1}^{N}FROM_{it}(d) \tag{19}$$

Baruník and Křehlík [44] argue that the above formula only describes fluctuations within specific frequency bands. The contribution of the given frequency band d to the total spillover can be measured by a weighted metric (from Equations 20–24), where the weight is $\Gamma(d) = \Sigma_{i,j=1}^{N}\tilde{\theta}_{ij,t}(d)/N$.

$$NPDC_{ij,t}(d)' = \Gamma(d)\cdot NPDC_{ij,t}(d) \tag{20}$$

$$TO_{it}(d)' = \Gamma(d)\cdot TO_{it}(d) \tag{21}$$

$$FROM_{it}(d)' = \Gamma(d)\cdot FROM_{it}(d) \tag{22}$$

$$NET_{it}(d)' = \Gamma(d)\cdot NET_{it}(d) \tag{23}$$

$$TCI_t(d)' = \Gamma(d)\cdot TCI_t(d) \tag{24}$$

The oil industry chain, encompassing upstream extraction, midstream processing, and downstream trade sectors, exhibits return that are influenced by a multitude of macroeconomic factors, policy interventions, and market supply-demand dynamics, with their effects evolving over time. The TVP-VAR-DY model, built upon the vector autoregression (VAR) framework and incorporating time-varying parameter (TVP) and dynamic correlation (DY) features, provides an effective analytical tool for investigating return spillover effects and dynamic interdependencies among these sectors.

4 Empirical analysis

4.1 Data

According to the division of the industry chain, upstream industries in the oil industry chain engage in oil extraction and development, while downstream industries engage in oil transportation, refining, processing, and sales. The oil industry stock index is one of the comprehensive indicators for measuring the actual operational performance of oil enterprises. Therefore, this article selects the stock indices of oil extraction (Ex), oil processing (Pr), and oil trade (Tr) industries from the tertiary industry index of Shenwan as representatives for studying the upstream, midstream, and downstream industries of oil, aiming to deeply explore the dynamic correlation and mutual influence between the stock index returns of the oil industry chain. We collect daily data spanning from 2 January 2014 to 26 April 2024, all source from the Eastmoney Choice Financial Terminal[1]. After screening and matching, a total of 2,496 samples are finally obtained. To calculate the return, we perform logarithmic first-order differencing on all data: $r_t = (\ln P_t - \ln P_{t-1}) \times 100$. Among them, r_t represents the return on day t and P_t represents the stock index price on day t.

Table 1 shows the descriptive statistics of each variable. According to the skewness coefficient, the stock index of the oil extraction industry shows a right-skewed feature, while the stock indices of both the oil processing and oil trade industries display left-skewed features; From the kurtosis perspective, the kurtosis of each variable is greater than 3, showing a hyperkurtosis distribution,

1 The data are available at https://choice.eastmoney.com/.

TABLE 1 Descriptive statistics.

Industry	Mean	Maximum	Minimum	Std. Dev	Skewness	Kurtosis	ADF
Ex	0.023	9.581	−10.524	1.732	0.088	10.211	−12.916[a]
Pr	−0.030	8.931	−10.269	1.557	−0.471	9.559	−10.100[a]
Tr	0.012	9.516	−11.596	2.230	−0.820	7.058	−11.821[a]

Note: This table reports descriptive statistics on the returns of oil extraction (Ex), oil processing (Pr), and oil trade (Tr) industries indices.
[a]Indicates that the representation is significant at the significance level of 1%, and ADF is an Augmented Dickey-Fuller unit root test statistic.

TABLE 2 Total connectedness.

Industry	Ex	Pr	Tr	From
Ex	53.00	34.86	12.14	47.00
Pr	33.94	51.40	14.66	48.60
Tr	14.78	18.46	66.75	33.25
To	48.72	53.33	26.80	
Net	1.72	4.73	−6.44	TCI = 42.95

Note: This table reports the total spillover matrix for the oil extraction (Ex), oil processing (Pr), and oil trading (Tr) industry indices. The elements in the *i*th row and *j*th column represent the contribution of market *j*'s price changes to market *i*'s price returns, representing the spillover of returns from market *j* to market *i*. The diagonal elements in the matrix quantify the shares attributed to each market's own price changes, whereas the non-diagonal elements represent cross-market spillover effects, i.e., spillovers from other markets. "To" and "From" denote the total directional spillovers, with "To" representing the total spillovers propagated to other markets (propagation) and "From" representing the total spillovers received from other markets (acceptance), respectively. These are calculated as the sums of the non-diagonal elements in the rows and columns, respectively. "Net" refers to the net spillover, which is the result of subtracting the total received ("From") from the total propagated ("To"). The total spillover effect index, shown in the bottom right corner of the table, is obtained by dividing the sum of all non-diagonal elements by the number of research variables.

TABLE 3 Frequency connectedness.

Industry	Ex	Pr	Tr	From
Panel A: Short-term frequency(1–5 days)				
Ex	42.77	28.00	9.64	37.64
Pr	27.88	41.71	11.59	39.47
Tr	12.28	14.68	53.26	26.97
To	40.16	42.68	21.24	
Net	2.52	3.21	−5.73	TCI = 34.69
Panel B: Medium-term frequency(5–22 days)				
Ex	7.48	4.93	1.90	6.83
Pr	4.76	7.33	2.43	7.19
Tr	1.97	2.78	10.09	4.75
To	6.72	7.71	4.34	
Net	−0.11	0.52	−0.41	TCI = 6.26
Panel C: Long-term frequency(22-Inf)				
Ex	2.74	1.93	0.60	2.53
Pr	1.30	2.36	0.63	1,94
Tr	0.53	1.00	3.40	1.53
To	1.83	2.93	1.23	
Net	−0.70	1.00	−0.30	TCI = 2.00

Note: The table presents the spillover results of stock indices in the oil extraction (Ex), oil processing (Pr), and oil trade (Tr) industries across different frequency domains. The elements in the *i*th row and *j*th column represent the contribution of market *j*'s price returns to market *i*'s price returns, specifically referring to the directional spillover of returns from market *j* to market *i*. The diagonal elements in the matrix quantify the shares attributed to the intrinsic volatility of each market, whereas the non-diagonal elements represent cross-market spillovers, i.e., spillover effects originating from other markets. "To" and "From" signify the total outgoing and incoming directional spillovers, respectively, calculated as the sums of the non-diagonal elements in the rows (transmission) and columns (reception). "Net" refers to the net spillover, which is the difference between the total outgoing ("To") and incoming ("From") spillovers. The total spillover effect index, located in the bottom right corner of the table, is derived by dividing the sum of all non-diagonal elements by the number of research variables.

which suggests a relatively concentrated distribution of data. After ADF testing, all sequences significantly rejected the null hypothesis of the existence of unit roots, thus proving that all variables are stationary. Therefore, we can proceed with modeling and analysis.

4.2 Static analysis

To analyze the connectivity of the total spillover of stock index returns in the oil extraction, oil processing, and oil trade industries in the time domain and short-term (1–5 days), medium-term (5–22 days), and long-term (22-infinite extension) frequencies, we calculate connectivity matrices in both the time and frequency domains. During the modeling process, we selected the optimal lag length of the VAR model as the 7th order based on the AIC criterion, and the results are presented in Tables 2, 3.

Table 2 reports the total spillover matrix of stock indices in the oil extraction, oil processing, and oil trading industries.

The data in Table 2 reveal that the diagonal elements of the matrix generally exhibit higher values than the non-diagonal elements, indicating that each variable has the most significant spillover effect on itself. In other words, the impact of the returns of each variable on itself dominates. The total spillover effect in the oil extraction, oil processing, and oil trading industries is as high as 42.95%, indicating that nearly half of the return variation among the three markets can be attributed to the impacts from the other two markets, which means a significant spillover relationship among variables. From the Net Spillover Index (Net), it can be seen that both the oil extraction industry and the oil processing industry have a net spillover greater than 0, indicating that they are the transmitters of spillover in the research system. Conversely, the oil trading industry has a net spillover less than 0, indicating that it is the recipient of spillovers.

Table 3 presents a detailed list of the changes in spillover effects in different frequency domains. We observe that the spillover effects between the three markets were 34.69%, 6.26%, and 2.00% in the short, medium, and long-term frequency domains, respectively. Specifically, the BK model divides the total spillover into three-time spans and we study the different degrees of spillover in these three time spans. The frequency domain results in Table 3 show that as the time span increases, the spillover index among the three markets gradually diminishes. In addition, we also found that regardless of whether in short, medium, or long cycles, the oil trading industry has always been a transmitter of return spillovers, the oil processing industry has always been a receiver of spillovers, and the oil extraction industry transitions from being a transmitter to a risk receiver.

4.3 Dynamic analysis

Under the influence of special events, the market price of the crude oil industry chain may undergoes significant fluctuations. Static spillover analysis mainly focuses on evaluating the average impact of each variable across the entire sample period. However, this approach has inherent limitations, as it struggles to effectively capture the time-varying nature of spillovers between markets and the specific spillover effects within distinct frequency domains. Therefore, we further analyze the dynamic spillover effects between markets from the perspective of each frequency domain. Figure 1 provides a visual display of three frequency domain spillover indices: short-term, medium-term, and long-term, where the total spillover across the time domain is equal to the sum of the spillovers in each frequency domain.

As shown in Figure 1, the overall trend reveals significant volatility in the total spillover index of the system. The fluctuation range of the total return spillover index between markets generally falls between 20% and 50%, and in exceptional cases, the peak can reach approximately 60%, indicating that the spillover effects between the oil extraction, oil processing, and oil trading industries exhibit time-varying characteristics. Notably, the spillover effect primarily manifests in the short term, with impacts lasting less than a week, suggesting a rapid spread among markets. In contrast, the long-term spillover effects appear relatively weak, possibly attributed to the fact that the long-term behavior of the oil extraction, processing, and trading industries within the oil industry chain is primarily governed by their inherent fundamentals.

According to the time-domain and frequency-domain charts, there are five significant increases in the total spillover index, namely, the oil price crash in 2014, the Chinese stock market crash in 2015, the United Kingdom Brexit in 2016, the outbreak of the pandemic in 2020, and the Russia-Ukraine conflict in 2022. In 2014, amid weak demand for crude oil, the Organization of the Petroleum Exporting Countries (OPEC) continued to increase crude oil production in order to maintain its market share and engage in fierce competition with US shale oil companies. This move led to a continuous increase in oil supply and a rebalancing of supply and demand in the crude oil market, resulting in significant market volatility. This change has also intensified the spillover effects between relevant markets in the Chinese oil industry chain, making the entire industry chain face more complex and volatile risks. The "6.26" stock market crash in China was the main reason for the rise in the total spillover index in 2015. This crash caused significant turbulence in the Chinese oil market, with various industries facing strong external shocks, thereby strengthening the correlation between different markets. At the beginning of 2016, the total spillover index began to show an upward trend again, which may be attributed to two factors: one is Brexit, and the other is major political, economic, and terrorist events, including the oil crisis. In 2020, the global epidemic broke out, and as the spread of the disease continued to expand, related risks also intensified. The total spillover effect at different time periods sharply increased and reached a peak. The increase in uncertainty during this period has led to a more uniform response from investors to shocks, thereby enhancing the connections among markets at different time frequencies. In 2022, the Russia-Ukraine conflict again triggered market turbulence, leading to an increase in the total spillover effect in oil-related industries. With the implementation of a series of measures by the Chinese government aimed at restoring the economy and improving people's livelihoods, such as the temporary zero tax rate on coal imports in response to the Russia-Ukraine conflict, the exemption of VAT for small-scale taxpayers, and the normalization of epidemic control measures, the spillover effect is gradually declining. At this point, the total spillover index began to stabilize, though it still remains relatively high compared to pre-extreme-event levels. In summary, whenever the socio-economic situation confronts major events or extreme risks, the spillover effects among oil-related industry markets will significantly intensify. As the economy and society gradually enter a period of stable development, this risk effect will also gradually stabilize.

Based on the investigation of the time-varying spillover effects of total returns in multiple markets, we further conduct in-depth analysis of the specific time-varying spillover effects between each market. Figures 2, 3 show the spillover effects and inflow effects[2] of each market in the sample in the time domain and different frequency domains.

From Figures 2, 3, we find that first, the directional spillover index reflects the time-varying nature of spillover effects between the oil extraction, oil processing, and oil trade industries. Secondly, from the results of directional spillover indices, the spillover indices and inflow indices of various markets show certain similarities in overall trends, with their changes mainly concentrated between 1%

2 The spillover index serves as an indicator of the degree to which a market is influenced by other markets, while the inflow index reveals the degree to which a market is influenced by other markets.

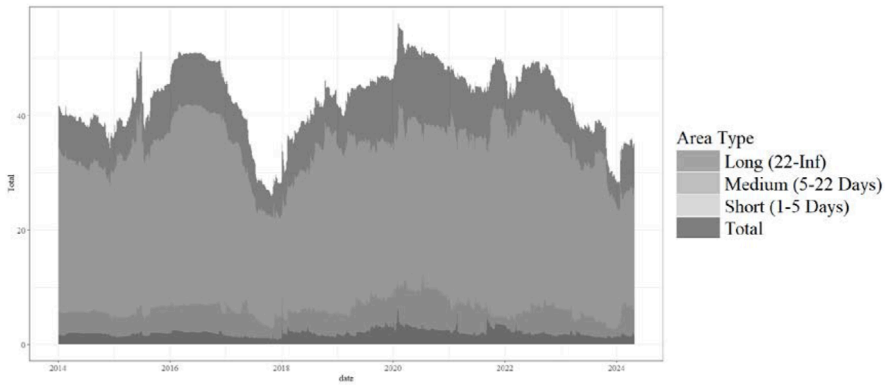


FIGURE 1
Dynamic total connectedness in various frequencies. This figure presents the trend in dynamic total spillover effects across different frequency bands within Chinese oil industry chain, spanning from 2 January 2014, to 26 April 2024. The analysis is based on the results derived from the TVP-VAR-DY model and the BK model. The frequency bands considered include the short-term band (set at 1–5 days, reflecting the 5-day trading week in stock markets), the medium-term green band (spanning 5–22 days, considering approximately 22 trading days in a month), and the long-term band (exceeding 22 days). The total spillover effect is calculated as the summation of spillover effects across these three frequency bands: short-term, medium-term, and long-term.

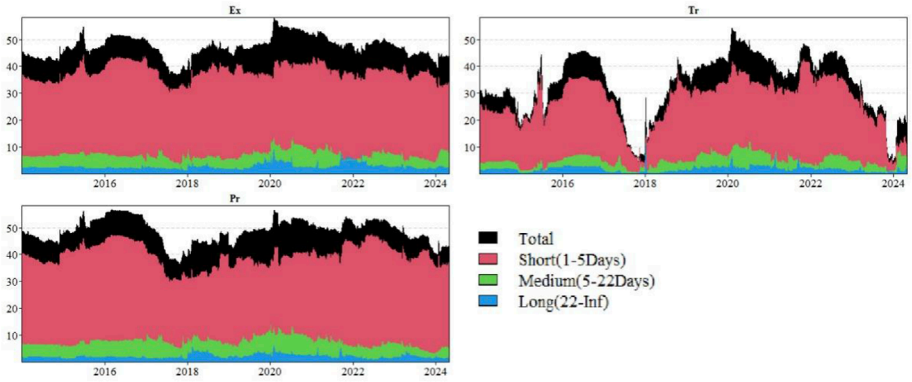


FIGURE 2
Dynamic total directional connectedness to others. This figure illustrates the trend in spillover effects across various frequency domains for different markets within Chinese oil industry chain, spanning from 2 January 2014, to 26 April 2024. The analysis is based on the results derived from the TVP-VAR-DY model and the BK model. The red area depicts the spillover in the short-term frequency band (1–5 days), the green area represents the spillover in the medium-term frequency band (5–22 days), and the blue area reflects the spillover in the long-term frequency band (exceeding 22 days). The black area, on the other hand, presents the summation of spillover effects across these three frequency bands: short-term, medium-term, and long-term.

and 40%. Finally, the net spillover effects of all variables in the time domain and short-term frequency bands are very obvious, while the spillover effects in the medium and long-term frequency bands are relatively small, which is consistent with the results of the static analysis previously mentioned.

Due to the bidirectional nature of spillover effects, a higher spillover index only indicates stronger correlation between markets, but cannot accurately reveal the direction of net spillover from one market to another. Therefore, it is necessary to conduct in-depth analysis of the net spillover index to determine the direction of return spillovers in the oil extraction, oil processing, and oil trade industries.

According to the net spillover index graph in Figure 4, the net spillover index fluctuates between positive and negative values, showcasing its temporal variability. Predominantly, however, the oil extraction and processing market exhibits a positive trend in its net spillover index, indicating that these industries have a greater spillover effect on the oil trading industry than they receive from other markets. Conversely, the oil trading industry's net spillover index registers a negative value, revealing its primary role as a recipient of spillovers. These findings suggest that within the oil industry chain, the extraction sector, as the primary source of crude oil resources, directly influences market supply through its production decisions and output adjustments. The processing sector, relying on crude oil supplied by extraction, transforms raw materials into various petroleum products through advanced refining technologies. Both sectors occupy relatively upstream positions in the value chain. Due to their control over critical resources and processing capabilities, they exert greater dominance in pricing and supply dynamics, more actively transmitting market

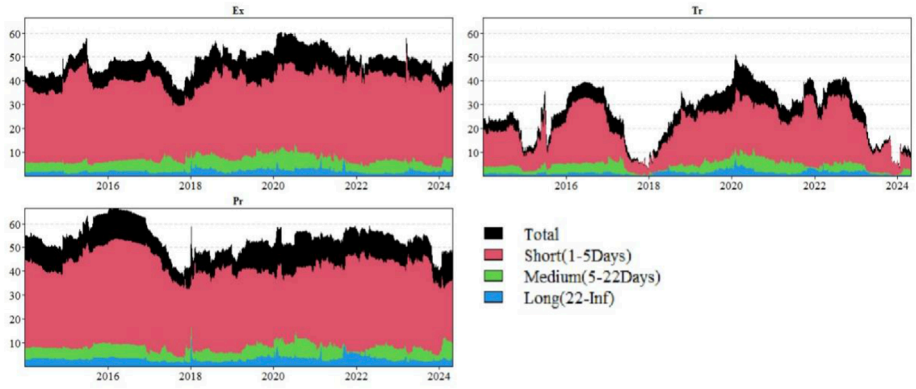


FIGURE 3
Dynamic total directional connectedness from others. This figure presents the trend in inflow spillover effects across different frequency domains for various markets within Chinese oil industry chain, spanning from 2 January 2014, to 26 April 2024. The analysis is based on the results derived from the TVP-VAR-DY model and the BK model. The red area illustrates the inflow spillover in the short-term frequency band (1–5 days), the green area represents the inflow spillover in the medium-term frequency band (5–22 days), and the blue area depicts the inflow spillover in the long-term frequency band (exceeding 22 days). The black area, in contrast, shows the aggregated result of inflow spillover effects across these three frequency bands: short-term, medium-term, and long-term.

fluctuations to the trade sector while receiving comparatively limited spillover effects from other markets. Conversely, the oil trade sector operates at the mid-to-downstream level of the industrial chain. In terms of market structure, it faces constraints from upstream supply conditions (extraction and processing) while simultaneously responding to demand fluctuations from end consumers. Additionally, external factors such as trade policies, tariff adjustments, and import-export quotas significantly impact its operations. These combined factors render the trade sector more susceptible to external shocks, primarily absorbing spillover effects transmitted from upstream industries rather than generating substantial outward spillovers. From this analysis, it becomes evident that while directional spillover effects can illuminate the overall beneficial impact of a variable on others within the system, the degrees of correlation and sensitivity among variables vary. Consequently, further analysis and discussion are paramount.

To delve into the dynamic mechanism of mutual influence between two industries, we employed a spillover index model with time-varying parameters, analyzing the pairwise spillover effects of returns across industries. The results are depicted in Figure 5.

From Figure 5, we find that the net paired spillover index of the oil extraction and processing industries varies significantly between positive and negative values, revealing the possibility of bidirectional time-varying asymmetric spillover effects between markets. The size and direction of this net spillover index often fluctuate significantly at certain critical moments, driven by multiple factors. Specifically, the net spillover index of the extraction industry initially showed a significant positive value, followed by a significant change in 2016, which was likely closely related to major political and economic events at the time, such as Brexit. After 2019, the index rose significantly again and became positive, which may be closely related to multiple factors such as energy price fluctuations, the pandemic, and international political events (such as the Russia-Ukraine conflict). These factors work together in the oil market, leading to increased uncertainty in oil supply and costs, which in turn affects downstream industries. In the oil processing industry market, the positive and negative fluctuations of spillover effects reflect their high sensitivity to market dynamics in the extraction industry. During the sample period, the oil trading industry's net spillover index mostly showed negative values, indicating that the industry played a more significant role as a spillover receiver for most of the time.

To further investigate the directional spillover effects of different variables, we constructed the spillover network diagram, which is shown in Figure 6. From left to right, we have the DY time domain, short-term, medium-term, and long-term frequency domains. Blue represents the net transmitter of spillover, and yellow represents the net receiver of spillover. The arrows in the figure visually display the direction of the spillover effect, whereas the thickness of the arrows accurately reflects the strength of the spillover effect. Specifically, the thicker the arrow, the more significant the corresponding spillover effect, that is, the greater the impact of one market or variable on another market or variable. This representation allows us to intuitively understand the degree of mutual influence between different markets or variables.

As depicted in Figure 6, the oil processing industry exhibits spillover effects on both oil extraction and oil trade, and the oil trade industry is mainly affected by spillover effects. This observation can be attributed to the oil processing industry's pivotal position in the industrial chain, bridging upstream oil extraction and downstream oil trade and terminal consumption. Consequently, its operational status carries significant weight in the industry chain, directly impacting the operations and profitability of both upstream and downstream enterprises. On the other hand, the oil trading industry is more susceptible to fluctuations in the global oil supply-demand relationship. In instances of global oversupply or inadequate demand for oil, oil prices tend to decline, subsequently compressing profitability within the oil trading industry. This finding aligns with the aforementioned analysis.

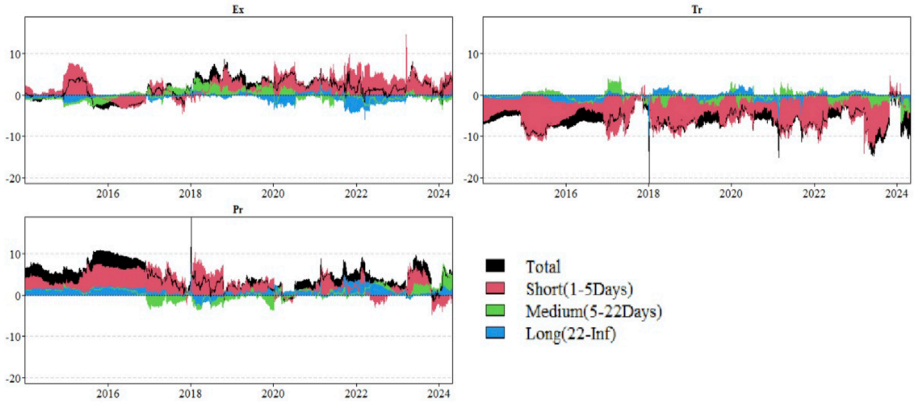


FIGURE 4
Dynamic net directional connectedness. This figure presents an analysis based on the TVP-VAR-DY model and the BK model. It reveals the dynamic evolution of net spillover effects across various markets within Chinese oil industry chain across different frequency domains, spanning from 2 January 2014, to 26 April 2024. These net spillover effect indices are derived by calculating the difference between the "outward spillovers" (To) and "inward spillovers" (From) for each market variable. Positive values indicate that the market variable acts as a net transmitter to other variables across the entire industry chain, whereas negative values suggest that the market variable is a net receiver, primarily receiving influences or information from other variables. The red areas reflect net spillovers in the short-term frequency band (1–5 days); green areas represent net spillovers in the medium-term frequency band (5–22 days); and blue areas depict net spillovers in the long-term frequency band (beyond 22 days). The black areas present the summation of net spillover effects across the short-term, medium-term, and long-term frequency bands.

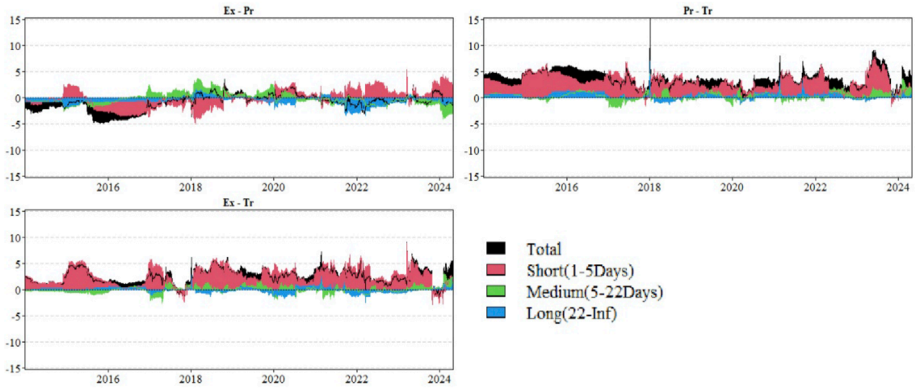


FIGURE 5
Dynamic net-pairwise directional connectedness. This figure presents the trend in net pairwise spillover effects across various markets within Chinese oil industry chain, spanning from 2 January 2014, to 26 April 2024. The analysis is based on the TVP-VAR-DY model and the BK model. The red area depicts the net pairwise spillovers in the short-term frequency band (1–5 days); the green area represents those in the medium-term band (5–22 days); and the blue area indicates the net pairwise spillovers in the long-term band (beyond 22 days). The black area, on the other hand, summarizes the net pairwise spillover effects across all three frequency bands—short-term, medium-term, and long-term.

4.4 Robustness test

To verify whether the empirical findings of this paper are overly reliant on the choice of model parameters, we adopt the robustness check methodology employ by Pavlova et al. [46], Wu et al. [38], Kočenda and Moravcová [47], among others. By adjusting the lag order of the TVP-VAR model, we investigate the sensitivity of spillover effects to variations in the lag order. On the basis of the benchmark 7th order, we additionally construct TVP-VAR models with 5th, 6th, 8th, and 9th orders. Subsequently, based on the estimation outcomes of these models featuring varying lag orders, we calculate the corresponding total spillover index. The results are shown in Figure 7, which shows the variation of the total spillover index for lag orders of 5th, 6th, 8th, and 9th from top to bottom.

From the perspective of the size of the total spillover index, both Figures 1, 7 fluctuate roughly within the range of 20%–50%, with similar values, indicating a significant spillover effect in the returns of Chinese oil extraction industry, oil processing industry, and oil trade industry. In terms of the trend in the total spillover index's variations, both Figures 1, 7 show obvious time-varying characteristics, and the trend over time is basically consistent. Especially in extreme situations, the total spillover index shows a significant increase, confirming that the total spillover index of Chinese oil extraction, oil processing, and oil trading markets is highly sensitive to extreme economic events. Generally, the total spillover index in Figures 1, 7 has experienced five periods of significant growth, namely, the period of oil price collapse in 2014, the period of Chinese stock market crash in 2015, the period of

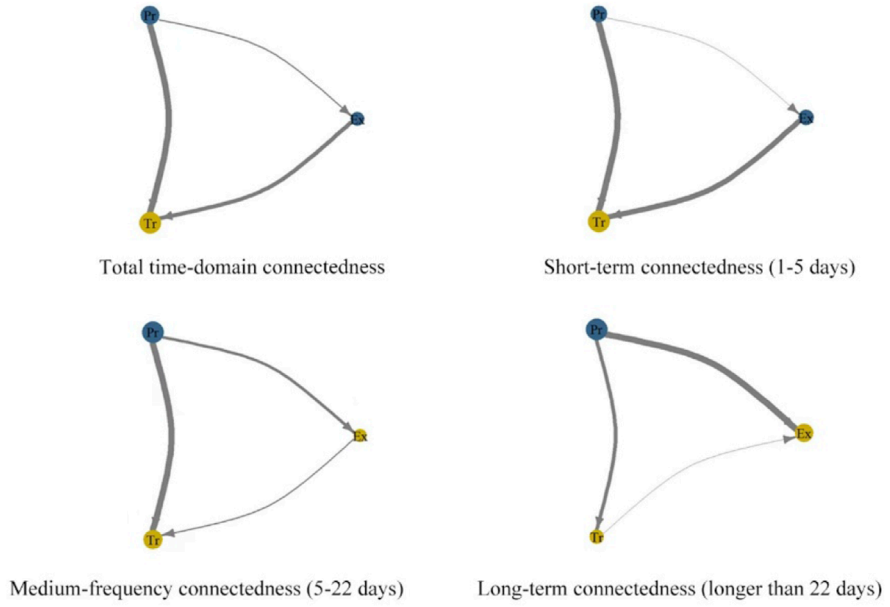


FIGURE 6
Directional connectedness network. This figure presents a visual representation where each edge connecting two nodes signifies the net pairwise spillover between two markets. The direction of the arrow indicates which market receives the shock from which market, and the thickness of the arrow reflects the intensity of the spillover effect. Blue nodes represent markets that are the originators of shocks within the system, while yellow nodes signify markets that are the recipients of shocks within the system.

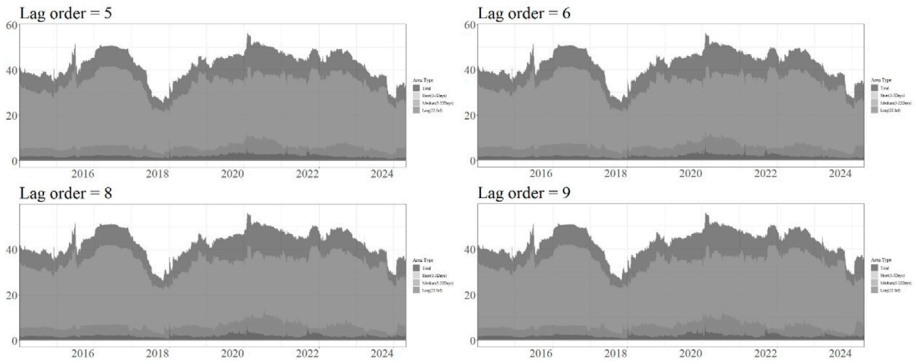


FIGURE 7
Dynamic total connectedness with different lag orders. This figure presents the trend of dynamic total spillover effects across different frequency domains for China's oil industry chain, utilizing the TVP-VAR model with lag orders of 5, 6, 8, and 9, respectively.

Brexit in 2016, the period of infectious disease outbreak in 2020, and the period of Russia-Ukraine conflict in 2022. In summary, the empirical results of this article are robust and not overly dependent on the lag order chosen by the model.

5 Concluding remarks

5.1 Conclusion

This paper employs the Time-Varying Parameter Generalized Variance Decomposition Spillover Index Model (TVP-VAR-DY) and the Baruník and Křehlík (BK) model, combined with a time-frequency domain perspective, to conduct an in-depth analysis of both static and dynamic spillover effects within China's oil industry chain system. The research not only calculates static spillover indices to evaluate the overall spillover intensity and direction within the system but also obtains dynamic spillover indices through a rolling sample window approach. This allows for the examination of spillover fluctuations throughout the entire sample period. Furthermore, by incorporating significant events, the study explores the changes in price linkage and risk contagion within the system in response to external shocks. Firstly, the results reveal significant dynamic and asymmetric spillover effects among

the three major industries of oil extraction, oil processing, and oil trade in the Chinese oil market. This finding reveals that price fluctuations among these three industries are not isolated but intertwined and mutually influential. Specifically, the oil processing industry plays a major role as a net spillover source, with its price fluctuations having a significant impact on the oil extraction and oil trade industries. In contrast, the oil trade industry primarily acts as a net spillover receiver, with its price fluctuations being significantly influenced by the other two industries. This asymmetric spillover effect demonstrates the complex interaction among various segments within the oil industry chain. Secondly, a deeper analysis from the frequency dimension reveals significant differences in spillover effects at different frequencies. As the frequency increases, the spillover index among the three markets gradually decreases, indicating the time-varying nature of the spillovers among market returns. Specifically, in the short term, the spillover effect among markets is most significant, reaching 34.69%; in the medium term, the spillover effect weakens to 6.26%; and in the long term, the spillover effect further decreases to 2%. This finding reveals the strong mutual influence of market returns spillovers in the short term and their gradual weakening trend in the long term. Thirdly, when confronted with the impact of extreme events, the overall total spillover level of the system tends to significantly amplify. Different types of extreme events have significant impacts on the directional spillover effects of market returns and the net pairwise spillover effects among markets, exhibiting their respective differences and dynamic changes. This finding emphasizes the important influence of extreme events on the spillover effects within the oil industry chain.

5.2 Implications

Building upon these findings, several insights are proposed. Given the dynamic, bidirectional, and asymmetric spillover effects among Chinese petroleum extraction, processing, and trade industries, it is imperative to strengthen collaborative management across the upstream and downstream segments of the industrial chain. Enterprises in these sectors should establish a digital information-sharing platform to exchange real-time data on production, inventory, and sales, jointly develop risk prevention and control plans, and achieve efficient resource allocation and risk sharing.

As the oil processing industry serves as a critical net spillover source, it is essential to formulate precise risk management policies. Enterprises should adopt advanced risk assessment models to conduct quantitative analyses of risks such as raw material price fluctuations and market supply-demand imbalances. Additionally, they should utilize futures hedging to lock in costs and enhance risk resilience. Government regulatory agencies must develop a comprehensive market monitoring system, leveraging big data and blockchain technologies to track global oil market dynamics in real time.

Given the time-varying nature of inter-market spillover effects, short-term measures should include establishing market liquidity monitoring indicators to closely observe significant market fluctuations. Enterprises should flexibly adjust inventory and production schedules, while governments should promptly implement temporary regulatory policies. In the medium term, enterprises ought to optimize their investment portfolios based on market trends, and governments should refine energy strategic reserve mechanisms. In the long run, even if spillover effects diminish, continuous optimization of risk management mechanisms remains crucial. Governments should promote energy structure transition, and enterprises should explore new energy ventures.

In the face of extreme events, systemic spillover levels tend to amplify. Therefore, governments and enterprises must enhance preparedness by establishing comprehensive risk early-warning and emergency response mechanisms, improving extreme event response capabilities, and proactively mitigating potential crises. Furthermore, international cooperation and exchange should be strengthened to collectively address the challenges and risks confronting the global oil market.

5.3 Limitations and future works

This study also acknowledges certain limitations. While it recognizes the impact of extreme events on the overall system spillover level, it does not delve deeply into the distinctions among different types of extreme events, such as natural disasters, policy shifts, and fluctuations in international oil prices, or their specific impact mechanisms. Furthermore, the primary focus of this study is on the return spillover effects within the oil industry chain, without considering the influence of external factors, such as the global economic environment and geopolitical risks, on these internal spillover effects. Future research could consider utilizing advanced models such as high-frequency data models or machine learning algorithms for comparative validation, refining the classification of extreme events and analyzing their impact mechanisms, as well as expanding the analysis to include external linkages of the industry chain to gain a more comprehensive understanding of the dynamic changes within the oil industry chain.

Data availability statement

Data will be made available on request.

Author contributions

LL: Supervision, Writing – original draft. YW: Writing – original draft, Writing – review and editing, Formal Analysis. LW: Writing – review and editing. YM: Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Guizhou Province Major Science and Technology Achievement Transformation Project [QKHCG(2024)ZD016], Guizhou Provincial Education

Department Natural Science Research Project [Qian Jiao Ji (2023) No. 033]; Key Project of Education Science Planning in Guizhou Province "Research on Evaluation and Improvement of Layered Blended Teaching Effect Based on MOOC" (Project No. 2020A046).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Liang C, Wei Y, Li X, Zhang X, Zhang Y. Uncertainty and crude oil market volatility: new evidence. *Appl Econ* (2020) 52(27):2945–59. doi:10.1080/00036846.2019.1696943

2. Jie LI, Huang L, Ping LI. Are Chinese crude oil futures good hedging tools? *Finance Res Lett* (2021) 38:101514. doi:10.1016/j.frl.2020.101514

3. Yang Y, Ma YR, Hu M, Zhang D, Ji Q. Extreme risk spillover between Chinese and global crude oil futures. *Finance Res Lett* (2021) 40:101743. doi:10.1016/j.frl.2020.101743

4. Liang C, Liao Y, Ma F, Zhu B. United States Oil Fund volatility prediction: the roles of leverage effect and jumps. *Empirical Econ* (2022) 62(5):2239–62. doi:10.1007/s00181-021-02093-5

5. Wei Y, Wang Z, Li D, Chen X. Can infectious disease pandemic impact the long-term volatility and correlation of gold and crude oil markets? *Finance Res Lett* (2022) 47:102648. doi:10.1016/j.frl.2021.102648

6. Ma F, Lu X, Zhu B. Uncertainty and fluctuation in crude oil price: evidence from machine learning models. *Ann Operations Res* (2023) 345:725–55. doi:10.1007/s10479-023-05463-7

7. Nasir MA, Naidoo L, Shahbaz M, Amoo N. Implications of oil prices shocks for the major emerging economies: a comparative analysis of BRICS. *Energy Econ* (2018) 76:76–88. doi:10.1016/j.eneco.2018.09.023

8. Abuzayed B, Al-Fayoumi N. Risk spillover from crude oil prices to GCC stock market returns: new evidence during the COVID-19 outbreak. *The North Am J Econ Finance* (2021) 58:101476. doi:10.1016/j.najef.2021.101476

9. Chen W, Huang Z, Yi Y. Is there a structural change in the persistence of WTI–Brent oil price spreads in the post-2010 period? *Econ Model* (2015) 50:64–71. doi:10.1016/j.econmod.2015.06.007

10. Elder J, Miao H, Ramchander S. Price discovery in crude oil futures. *Energy Econ* (2014) 46:S18–S27. doi:10.1016/j.eneco.2014.09.012

11. Liao HC, Lin SC, Huang HC. Are crude oil markets globalized or regionalized? Evidence from WTI and Brent. *Appl Econ Lett* (2014) 21(4):235–41. doi:10.1080/13504851.2013.851766

12. Caporin M, Fontini F, Talebbeydokhti E. Testing persistence of WTI and Brent long-run relationship after the shale oil supply shock. *Energy Econ* (2019) 79:21–31. doi:10.1016/j.eneco.2018.08.022

13. Hammoudeh S, Li H, Jeon B. Causality and volatility spillovers among petroleum prices of WTI, gasoline and heating oil in different locations. *The North Am J Econ Finance* (2003) 14(1):89–114. doi:10.1016/S1062-9408(02)00112-2

14. Ji Q, Fan Y. Dynamic integration of world oil prices: a reinvestigation of globalisation vs. regionalisation. *Appl Energy* (2015) 155:171–80. doi:10.1016/j.apenergy.2015.05.117

15. Leong SH. Global crude oil and the Chinese oil-intensive sectors: a comprehensive causality study. *Energy Econ* (2021) 103:105558. doi:10.1016/j.eneco.2021.105558

16. Huang BN, Yang CW, Hwang MJ. The dynamics of a nonlinear relationship between crude oil spot and futures prices: a multivariate threshold regression approach. *Energy Econ* (2009) 31(1):91–8. doi:10.1016/j.eneco.2008.08.002

17. Lei C, Yong Z. The properties and cointegration of oil spot and futures prices during financial crisis. *Energy Proced* (2011) 5:353–9. doi:10.1016/j.egypro.2011.03.060

18. Lee CC, Zeng JH. Revisiting the relationship between spot and futures oil prices: evidence from quantile cointegrating regression. *Energy Econ* (2011) 33(5):924–35. doi:10.1016/j.eneco.2011.02.012

19. Wang Y, Wu C. Are crude oil spot and futures prices cointegrated? Not always. *Econ Model* (2013) 33:641–50. doi:10.1016/j.econmod.2013.05.013

20. Mehrara M, Hamldar M. The relationship between spot and futures prices in Brent oil crude market. *Int Lett Social Humanistic Sci* (2014) 39:15–9. doi:10.18052/www.scipress.com/ilshs.39.15

21. Chen PF, Lee CC, Zeng JH. The relationship between spot and futures oil prices: do structural breaks matter? *Energy Econ* (2014) 43:206–17. doi:10.1016/j.eneco.2014.03.006

22. Chang CP, Lee CC. Do oil spot and futures prices move together? *Energy Econ* (2015) 50:379–90. doi:10.1016/j.eneco.2015.02.014

23. Klein T. Trends and contagion in WTI and Brent crude oil spot and futures markets-The role of OPEC in the last decade. *Energy Econ* (2018) 75:636–46. doi:10.1016/j.eneco.2018.09.013

24. Magkonis G, Tsouknidis DA. Dynamic spillover effects across petroleum spot and futures volatilities, trading volume and open interest. *Int Rev Financial Anal* (2017) 52:104–18. doi:10.1016/j.irfa.2017.05.005

25. Guo S, Li H, An H, Ma N, Sun Q, Feng S, et al. Detecting the horizontal/vertical price relationship patterns in the global oil industry chain through network analysis. *Energy* (2024) 296:131054. doi:10.1016/j.energy.2024.131054

26. Qi Y, Bai J, Liu S. Spillover dynamics among commodities along the Chinese oil industrial chain: from the perspective of multidimensional networks. *Int Rev Econ & Finance* (2024) 96:103612. doi:10.1016/j.iref.2024.103612

27. Huang X, Huang S. Identifying the comovement of price between China's and international crude oil futures: a time-frequency perspective. *Int Rev Financial Anal* (2020) 72:101562. doi:10.1016/j.irfa.2020.101562

28. Palao F, Pardo Á, Roig M. Is the leadership of the Brent-WTI threatened by China's new crude oil futures market? *J Asian Econ* (2020) 70:101237. doi:10.1016/j.asieco.2020.101237

29. Ji Q, Zhang D, Zhao Y. Intra-day co-movements of crude oil futures: China and the international benchmarks. *Ann Operations Res* (2022) 313(1):77–103. doi:10.1007/s10479-021-04097-x

30. Wang J, Qiu S, Yick HY. The influence of the Shanghai crude oil futures on the global and domestic oil markets. *Energy* (2022) 245:123271. doi:10.1016/j.energy.2022.123271

31. Lee CC, Zhou H, Xu C, Zhang X. Dynamic spillover effects among international crude oil markets from the time-frequency perspective. *Resour Policy* (2023) 80:103218. doi:10.1016/j.resourpol.2022.103218

32. Antonakakis N, Chatziantoniou I, Gabauer D. Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *J Risk Financial Management* (2020) 13(4):84. doi:10.3390/jrfm13040084

33. Diebold FX, Yilmaz K. Better to give than to receive: predictive directional measurement of volatility spillovers. *Int J Forecast* (2012) 28(1):57–66. doi:10.1016/j.ijforecast.2011.02.006

34. Diebold FX, Yılmaz K. On the network topology of variance decompositions: measuring the connectedness of financial firms. *J Econom* (2014) 182(1):119–34. doi:10.1016/j.jeconom.2014.04.012

35. Shah AA, Paul M, Bhanja N, Dar AB. Dynamics of connectedness across crude oil, precious metals and exchange rate: evidence from time and frequency domains. *Resour Policy* (2021) 73:102154. doi:10.1016/j.resourpol.2021.102154

36. Mo B, Meng J, Zheng L. Time and frequency dynamics of connectedness between cryptocurrencies and commodity markets. *Resour Policy* (2022) 77:102731. doi:10.1016/j.resourpol.2022.102731

37. Hoque ME, Soo-Wah L, Billah M. Time-frequency connectedness and spillover among carbon, climate, and energy futures: determinants and portfolio risk management implications. *Energy Econ* (2023) 127:107034. doi:10.1016/j.eneco.2023.107034

38. Wu H, Zhu H, Huang F, Mao W. How does economic policy uncertainty drive time–frequency connectedness across commodity and financial markets? *The North Am J Econ Finance* (2023) 64:101865. doi:10.1016/j.najef.2022.101865

39. Liang C, Goodell JW, Li X. Impacts of carbon market and climate policy uncertainties on financial and economic stability: evidence from connectedness network analysis. *J Int Financial Markets, Institutions Money* (2024) 92:101977. doi:10.1016/j.intfin.2024.101977

40. Polat O, Ertuğrul HM, Sakarya B, Akgül A. TVP-VAR based time and frequency domain food & energy commodities connectedness an analysis for financial/geopolitical turmoil episodes. *Appl Energy* (2024) 357:122487. doi:10.1016/j.apenergy.2023.122487

41. Koop G, Pesaran MH, Potter SM. Impulse response analysis in nonlinear multivariate models. *J Econom* (1996) 74(1):119–47. doi:10.1016/0304-4076(95)01753-4

42. Pesaran HH, Shin Y. Generalized impulse response analysis in linear multivariate models. *Econ Lett* (1998) 58(1):17–29. doi:10.1016/S0165-1765(97)00214-0

43. Chatziantoniou I, Gabauer D. EMU risk-synchronisation and financial fragility through the prism of dynamic connectedness. *The Q Rev Econ Finance* (2021) 79:1–14. doi:10.1016/j.qref.2020.12.003

44. Baruník J, Křehlík T. Measuring the frequency dynamics of financial connectedness and systemic risk. *J Financial Econom* (2018) 16(2):271–96. doi:10.1093/jjfinec/nby001

45. Stiassny A. A spectral decomposition for structural VAR models. *Empirical Econ* (1996) 21:535–55. doi:10.1007/BF01180700

46. Pavlova I, De Boyrie ME, Parhizgari AM. A dynamic spillover analysis of crude oil effects on the sovereign credit risk of exporting countries. *The Q Rev Econ Finance* (2018) 68:10–22. doi:10.1016/j.qref.2018.03.003

47. Kočenda E, Moravcová M. Frequency volatility connectedness and portfolio hedging of US energy commodities. *Res Int Business Finance* (2024) 69:102274. doi:10.1016/j.ribaf.2024.102274

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 23 June 2025
DOI 10.3389/fphy.2025.1548092

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Bing He,
Jiangsu Ocean University, China
Qianqian Zhang,
Beijing University of Civil Engineering and Architecture, China

*CORRESPONDENCE
Di Wu,
z5606663@ad.unsw.edu.au

RECEIVED 19 December 2024
ACCEPTED 28 May 2025
PUBLISHED 23 June 2025

CITATION
Wu D, An Q and Zheng L (2025) Does climate risk undermine urban economic resilience? —Evidence from Chinese cities.
Front. Phys. 13:1548092.
doi: 10.3389/fphy.2025.1548092

COPYRIGHT
© 2025 Wu, An and Zheng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Does climate risk undermine urban economic resilience? —Evidence from Chinese cities

Di Wu[1]*, Qiguang An[2] and Lin Zheng[2]

[1]School of Business, University of New South Wales, Kensington, NSW, Australia, [2]School of Statistics and Mathematics, Shandong University of Finance and Economics, Jinan, China

Climate change has become a critical global issue, influencing both natural environments and economic structures, especially in rapidly urbanizing regions. This paper examines the impact of climate risks on urban economic resilience, using panel data from Chinese cities between 2009 and 2022. A multidimensional indicator system, encompassing recovery resilience, adaptive resilience, and transformative innovation capacity, is developed to identify how climate risks negatively affect urban economic resilience. The results show that climate risks weaken resilience by reducing population size and destabilizing financial systems. Additionally, these risks have significant spatial spillover effects, extending from local areas to neighboring cities through regional economic networks, with particular impact on geographically adjacent cities. Heterogeneity analysis indicates that developed eastern cities, central cities, and resource-based cities are more vulnerable to climate risks, while cities in central and western regions and non-central cities show greater resilience. The paper proposes policy recommendations to strengthen urban resilience, including investment in climate-adaptive infrastructure, promoting economic diversification, establishing cross-regional climate risk management, developing green finance systems, and raising public awareness of climate risks.

KEYWORDS

climate risk, economic resilience, spatial spillover, population size, financial stability

1 Introduction

Since the 1990s, China has experienced rapid economic growth and has now become the world's second-largest economy. However, the rapid processes of industrialization and urbanization have also led to increasingly severe ecological issues and climate risks [1]. According to data from the United Nations, eight out of 54 major natural disasters worldwide in the 20th century occurred in China. In 2024, various natural disasters in China caused direct economic losses amounting to 401.11 billion RMB, highlighting the significant impact of climate risks on China's socio-economic development [2, 3, 4]. In particular, the "China Climate Change Blue Book (2024)," published in 2024, points out that China's climate risk index has continuously risen since 1961, with a notably high increase since the mid-1990s, aligning with the global trend of climate warming. In 2022, the Chinese government released the "National Climate Change Adaptation Strategy 2035," which explicitly proposes enhancing the capacity of society to cope with climate change and building an adaptive society by 2035. This indicates that climate risks are not only an ecological challenge but also a significant test for China's high-quality economic development, placing higher demands on the resilience of urban economic systems [3, 5]. In

recent years, the frequency and intensity of extreme weather events brought about by climate change have significantly increased, with their impact on urban economic systems being particularly notable. For example, extreme climate events such as heavy rainfall, heatwaves, and droughts have severely affected infrastructure and economic activities. According to the "Global Annual to Decadal Climate Update" released by the World Meteorological Organization (WMO) in 2024, climate change is approaching the risk limits set by the Paris Agreement. In the same year, national natural disasters affected 32.38 million people, with direct economic losses amounting to 93.16 billion yuan. These phenomena indicate that climate change has become a critical factor affecting the stability and sustainability of urban economies [2].

In the face of increasing external shocks and potential risks, the economic systems of different cities show significant differences. Some cities can quickly recover from shocks and maintain economic vitality, while others may suffer long-term damage, demonstrating weaker disaster resilience [6]. A key question arises: Has the increasing climate risk had a significant impact on urban economic systems? The term "resilience," initially a concept in physics, refers to a system's ability to return to its original state after a shock; in economics, it is extended to refer to the ability of an economic system to resist external shocks and achieve recovery, adaptation, and transformation [7, 8]. However, existing research has paid little attention to how climate risks specifically affect urban economic resilience and the mechanisms behind this, lacking quantitative analysis and empirical research, which makes it difficult to provide systematic references for practical policy design.

To fill this research gap, this paper explores the impact of climate risks on urban economic resilience based on panel data from Chinese cities between 2009 and 2022. The contributions of this paper are as follows: first, unlike previous studies that focus on a single dimension of economic resilience or macroeconomic impacts, this paper innovatively constructs a multidimensional economic resilience index system and reveals the multi-path impact of climate risks on urban economic resilience. Second, although existing literature has focused on the direct impact of climate risks on specific cities, few studies have explored the cross-city and regional spatial spillover effects. This paper quantifies the significant negative impact of climate risks on the economic resilience of neighboring cities through spatial Durbin models, considering the geographic, economic, and proximity dimensions. Finally, compared to studies focusing on a single region, this paper uses a broader city panel dataset to analyze the differentiated impact of climate risks on cities in different regions and types, providing a novel research perspective.

The structure of the paper is as follows: Section 2 presents a literature review on climate risks and economic resilience; Section 3 introduces the theoretical framework and research hypotheses; Section 4 describes the empirical model and research methods; Section 5 presents the empirical results and analysis; Section 6 provides conclusions and policy implications.

2 Literature review

2.1 Literature review on climate risk research

With the continuous increase in human activities, particularly the expansion of economic activities, the sustainability of the environment is facing profound impacts and addressing climate risks requires environmental regulation [5, 9]. This imbalanced development model has further triggered a series of global climate issues. As a future uncertainty brought about by climate change and its negative effects, climate risk has become an important research topic in the global academic community. Climate risks are primarily divided into two major categories: transition risk and physical risk. Transition risk refers to the long-term economic impacts arising from policy changes, technological innovations, and market sentiment fluctuations during the process of society transitioning to a low-carbon economy or sustainable development [10]. For example, Campiglio pointed out that adjustments in climate policies may lead to increased operational costs for businesses and a loss of asset value [11].

In contrast, physical risk refers to the short-term and direct economic losses caused by extreme climate events or natural disasters, which typically have more significant destructive impacts on the economic system [12]. Typical physical risks include sea-level rise, extreme heat, storms, and floods, all of which pose major challenges to economic infrastructure and supply chains. Given the immediate and tangible nature of these risks, physical risks are particularly relevant to urban economic systems, where the impacts are more direct and observable. In addition, the availability of data on physical risks, such as historical records of extreme weather events, enhances the ability to quantify and model their effects on urban economies. In the economic field, physical climate risks have profound effects on financial markets, business operations, and investment decisions. For instance, Hu and Borjigin examined the volatility of climate risk in the futures market, Lin and Wu explored the impact of physical risks on financial products, and Campiglio analyzed the effects of climate risk on asset management and portfolio optimization [11, 13, 14].

At the macroeconomic level, physical risks have had widespread and far-reaching impacts on agricultural economies [15], external debt levels, and international trade [2, 16, 17]. In addition, Carleton suggested that the frequent occurrence of extreme climate events could further exacerbate global economic instability, particularly in regions with more vulnerable economies [12]. Based on past studies, physical risks have more directly affected the stability of economic systems, and under the backdrop of frequent extreme weather events, their impact has been increasing. Therefore, this paper focuses on the study of physical risks and their impact on economic resilience, aiming to reveal how to enhance the adaptive and recovery capacities of economic systems when facing natural disasters and extreme climate events.

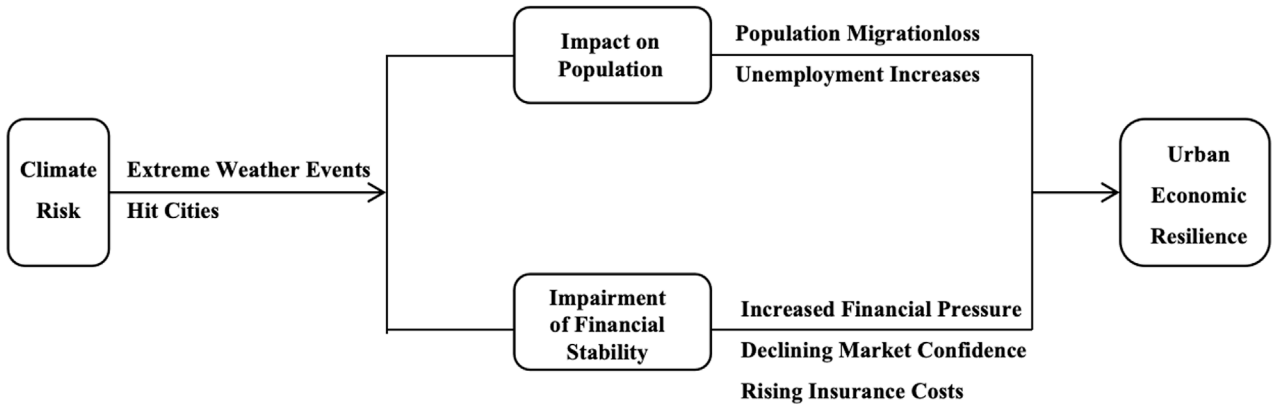


FIGURE 1
Pathway mechanisms of climate risk impact on urban economic resilience.

TABLE 1 Urban climate risk indicator system.

	Sub-indicators	Explanation
Climate risk index (*CRI*)	*LTD*	The number of extreme low-temperature days
	HTD	The number of extreme high-temperature days
	ERD	The number of extreme rainfall days
	EDD	The number of extreme drought days

Notes: Following the method of Guo et al. [38], we constructed an indicator system for urban climate risk.

2.2 Literature review on economic resilience research

The term "resilience" originates from physics and refers to the ability of materials to absorb energy during deformation and fracture. In economics, "economic resilience" is used to describe the ability of an economic system to resist and adjust its development path when faced with shocks [7, 8]. In past studies on the factors affecting economic resilience, both theoretical and empirical research have been emphasized. Theoretically, Sun and Sun reviewed relevant concepts and proposed research dimensions based on China's context [18]; Liu et al. further clarified the connotation of the economic resilience concept [19]. Empirically, scholars have explored factors influencing economic resilience from perspectives such as macroeconomics [8, 48], industrial diversity [20], smart supply chain [15] and digital finance [20]. However, most past research on economic resilience has focused on macroeconomics or specific industries [7, 8], with relatively limited analysis of the economic resilience of cities as complex dynamic systems. Furthermore, static analytical methods are often employed, making it difficult to fully capture the dynamic adjustment process of economic resilience in the context of multidimensional shocks [20].

Research on the impact of climate risk on economic resilience has found that physical risks affect economic resilience through various pathways. On one hand, extreme climate events directly damage infrastructure and production systems, weakening the short-term resilience of economic systems. For example, Carleton pointed out that natural disasters often interrupt supply chains and logistics systems, causing profound effects on economic activities [12]. On the other hand, physical risks affect economic resilience through indirect pathways, such as market fluctuations, declining investment confidence, and long-term policy adjustments. For instance, Ma et al. [10] emphasized that policy changes can lead to shifts in economic structures and long-term challenges for resilience. Moreover, Zhao et al. [21] explored the effects of climate change on urban economic resilience, Liu et al. [19] examined the role of smog in urban economic resilience, and Wu et al. [22] assessed the economic resilience performance of the Yellow River Basin under climate change. Additionally, Liu et al. [19] studied the negative impact of air pollution on urban economic resilience, emphasizing the importance of environmental governance in enhancing urban resilience.

While existing research has revealed some mechanisms by which climate risk impacts economic resilience, several gaps remain. First, the dynamic relationship between climate risk and economic resilience, particularly in the context of frequent extreme weather events, remains underexplored [23]. Second, systematic studies on the heterogeneous impacts of climate risks on urban economic resilience are lacking [9]. Finally, most current research has focused on the impact of physical climate risks on macroeconomics or specific industries, with limited exploration of the performance

TABLE 2 Urban economic resilience indicator system and weights.

Dimension and weight	Indicator explanation and units	Nature
Resistance and recovery capacity (0.178)	Per capita GDP (10,000 yuan) 0.278	+
	Per capita disposable income of urban residents (10,000 yuan) 0.126	+
	Household savings deposits (10,000 yuan) 0.401	+
	Average employee wage (yuan) 0.194	+
Adaptive and adjustment capacity (0.445)	Fixed asset investment (10,000 yuan) 0.421	+
	Local fiscal expenditure (10,000 yuan) 0.554	+
	Balance of loans and deposits in RMB from financial institution 0.025	+
Innovation and transformation capacity (0.377)	Total number of enterprises above designated size 0.164	+
	Education investment (10,000 yuan) 0.272	+
	Science and technology investment (10,000 yuan) 0.541	+
	Urbanization rate 0.023	+

Notes: In the indicator system for urban economic resilience, weights are calculated using the entropy weighting method.

of urban economic resilience, a complex dynamic system, under physical climate risk shocks. This paper aims to fill this research gap by combining theoretical analysis and empirical research to reveal the mechanisms through which physical climate risks impact urban economic resilience and to provide scientific evidence for policy-making.

3 Theoretical analysis and research hypotheses

As global climate change intensifies, climate risks not only profoundly affect the ecological environment but also pose a significant threat to the economic structure and stability of cities. Specifically, the impact of climate risks on key areas such as urban population size and the stability of financial markets is particularly prominent [4]. These areas, being fundamental components of urban economies, are directly linked to the core capacity of urban economic resilience. Therefore, this study constructs a framework for urban economic resilience based on recovery capacity, adaptive capacity, and transformative innovation capacity. This framework is developed to capture the multidimensional nature of resilience, addressing how different components interact to enhance a city's ability to respond to climate risks [24]. Existing literature indicates that climate risks primarily affect urban economic resilience through several pathways:

3.1 Direct impact mechanisms of climate risks on urban economic resilience

Currently, the direct impact of climate risks on cities is primarily manifested in the destruction of critical economic and infrastructure systems. Extreme weather events such as hurricanes, floods, and droughts significantly disrupt the normal operation of infrastructure systems, including transportation, energy supply, and water systems. These disruptions weaken the city's response and recovery capacities, delivering substantial shocks to economic productivity and residents' quality of life [25]. Specifically, interruptions in production activities may hinder business operations and lead to economic stagnation. Concurrently, business losses may increase unemployment rates and undermine the ability to meet basic living needs for residents, further exacerbating social instability. This social instability, in turn, diminishes the economic system's ability to recover from external shocks.

Additionally, the losses caused by climate risks will increase the financial burden on governments, forcing them to allocate large amounts of funds for post-disaster reconstruction and climate adaptation measures. This limits the flow of funds and resource allocation to other economic sectors. Such fiscal pressure not only affects urban economic resilience in the short term but may also erode the city's innovation capacity and economic growth potential in the long run [26]. Cities with relatively weak infrastructure may fall into prolonged economic distress, hindering their economic development and the achievement of sustainable development goals. This amplifies local fiscal pressures, forcing governments to cut budgets for key sectors such as education and healthcare, impacting long-term urban development and the economic resilience system [27]. Therefore, climate risks not only cause direct damage to the city's critical infrastructure but also generate profound effects on economic resilience and social stability through a series of cascading reactions.

Hypothesis 1: Climate risks negatively impact urban economic resilience through direct mechanisms.

TABLE 3 Definitions of key variables.

Variable category	Variable name	Abbreviation	Variable description	Data sources
Explanatory variable	Climate risk index	*CRI*	A composite index derived from the standardized indices of LTD. (Extreme Low Temperature Days), HTD (Extreme High Temperature Days), ERD (Extreme Rainfall Days), and EDD (Extreme Drought Days)	National meteorological centre of China
Dependent variable	Urban economic resilience index	*UERI*	Measured using three dimensions: resistance and recovery capacity, adaptive and adjustment capacity, and innovation and transformation capacity, with 11 secondary indicators	China city statistical yearbook
	Resistance and recovery capacity	*Rel*	Calculated using entropy weighting based on indicators such as *per capita* GDP and *per capita* disposable income of urban residents	China city statistical yearbook
	Adaptive and adjustment capacity	*Ada*	Calculated using entropy weighting based on indicators like fixed asset investment and local fiscal expenditure	China city statistical yearbook
	Innovation and transformation capacity	*Enpu*	Calculated using entropy weighting based on indicators like the total number of large-scale enterprises and fiscal education expenditure in the region	China city statistical yearbook
Mediating variable	Urban population size	*Psize*	Logarithmic value of the total urban population at the end of the year	China city statistical yearbook
	Urban financial stability	*Fin*	Ratio of year-end loan balances of financial institutions to regional GDP.	China city statistical yearbook
Control variable	Urban entrepreneurship activity	*Live*	Ratio of the number of private and individual employees in urban areas to the urban population	China city statistical yearbook
	Urban facility development level	*Fund*	Per capita road area	China city statistical yearbook
	Urban economic density	*Den*	Ratio of regional GDP to urban land area	China city statistical yearbook
	Urban foreign investment dependence	*Export*	Proportion of actual foreign investment in GDP for the city that year	China city statistical yearbook
	Urban human capital level	*Hr*	Ratio of regular college and university students to the permanent urban population	China city statistical yearbook

TABLE 4 Descriptive statistics of main variables.

Variable	N	Mean	Sd	Min	Med	Max
CRI	2,212	0.097	0.094	0.011	0.070	0.987
UERI	2,212	3.318	0.240	0.292	3.329	4.472
Rel	2,212	0.141	0.081	0.016	0.120	0.695
Ada	2,212	0.069	0.082	0.004	0.044	0.726
Enpu	2,212	0.041	0.065	0.002	0.023	0.794
Live	2,212	0.123	0.126	−0.023	0.089	0.130
Fund	2,212	2.505	0.678	0	2.577	4.112
Den	2,212	0.040	0.037	0.001	0.028	0.271
Export	2,212	0.017	0.018	0	0.011	0.136
Hr	2,212	0.018	0.018	0.0003	0.012	0.113

3.2 Pathways of climate risk impact on urban economic resilience

The impact of climate risks on urban economic resilience primarily occurs through two main pathways: reducing urban population size and disrupting urban financial stability. Specifically, extreme weather events weaken a city's economic resilience by disrupting various sectors, including population, infrastructure, and financial systems, which in turn affect its capacity to cope with and recover from external shocks. The specific mechanisms of these pathways are illustrated in Figure 1.

3.2.1 Impact on Urban Population and Labor Force

The economic vitality and resilience of cities largely depend on stable labor forces and a healthy population size. Densely populated areas tend to demonstrate stronger resilience, which positively influences regional economic resilience, and this relationship can vary depending on urban governance and infrastructure [28]. However, climate risks can reduce population size by altering mobility and living conditions, thus weakening urban economic resilience. Extreme weather events, for instance, damage urban living environments, compelling large numbers of residents to migrate, especially from frequently disaster-prone areas. Long-term adverse climate conditions force residents to relocate to safer regions [29]. Large-scale population outflows not only reduce labor supply but also shrink the consumer market, weakening the economic vitality of the city and increasing vulnerability during economic crises [30]. Furthermore, climate-related physical risks may disrupt businesses dependent on natural resources or logistics, leading to transportation system breakdowns and business shutdowns, which in turn drive up unemployment rates, reduce resident income, and shrink consumer spending [31]. By undermining population structure and labor markets, climate risks weaken urban economic resilience, making it more difficult for cities to demonstrate recovery capacity in the face of future external shocks.

3.2.2 Impact on Urban Financial Stability

Financial stability is a core factor in urban economic resilience, as it reduces economic volatility, supports innovation, and eliminates policy distortions, thereby significantly enhancing regional resilience [32]. However, climate risks increase uncertainty and vulnerability in financial systems and economic operations. Firstly, climate risks lower investor confidence, exacerbating fluctuations in financial markets. Companies facing operational interruptions and losses due to climate events may struggle to secure financial support, leading to financial crises or even bankruptcy. This credit tightening effect not only affects the impacted businesses but also spreads to the broader urban economy, making the financial system more fragile [33]. Moreover, the risks associated with climate events drive up insurance costs, leading to higher premiums. This increases operational costs for businesses, making insurance unaffordable for some enterprises and households, thus exposing them to greater risks and making it impossible to ensure asset security when facing future shocks. The instability in the insurance market also affects bank loans and debt markets, further weakening the city's economic resilience.

In summary, climate risks reduce urban economic resilience by decreasing population size and disrupting financial stability. Population declines lead to reduced economic vitality, labor market contraction, and increased unemployment, limiting the economy's ability to recover from external shocks. Fiscal pressure, financial volatility, and high insurance costs weaken the city's capacity to respond to future risks, further hindering economic stability and long-term development.

Hypothesis 2: Climate risks negatively affect urban economic resilience by reducing population size and disrupting financial stability.

3.3 Spatial spillover effects of the negative impact of climate risks on urban economic resilience

Climate risks have far-reaching impacts on the economic resilience of other cities through the transmission effects within economic networks. Extreme weather events such as hurricanes, floods, or droughts typically damage the infrastructure and economic activities of affected cities, and these impacts quickly spread through regional economic linkages. For example, supply chain disruptions may lead to shortages of raw materials and intermediate goods, hindering production in businesses reliant on affected cities. Logistics delays may disrupt interregional flows of goods, causing market supply-demand imbalances. The globalization of capital flows further exacerbates this transmission effect. When investors withdraw capital due to economic uncertainties in affected cities, capital outflows may destabilize the financial systems of other cities, thereby weakening their economic resilience [34, 35].

TABLE 5 Baseline regression.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CRI	*Rel*	*Ada*	*Enpu*	*CRI*	*Rel*	*Ada*	*Enpu*
UERI	−0.601**	−0.560	−0.455*	−0.373**	−0.575**	−0.519	−0.440*	−0.360**
	(0.302)	(0.415)	(0.259)	(0.183)	(0.254)	(0.390)	(0.225)	(0.148)
Live					0.134***	0.123***	0.094***	0.0935***
					(0.034)	(0.028)	(0.025)	(0.028)
Fund					0.007*	0.005*	0.004	0.006**
					(0.004)	(0.003)	(0.003)	(0.003)
Den					0.251	0.078	0.183	0.249
					(0.241)	(0.205)	(0.177)	(0.195)
Export					0.043	−0.142*	0.128	0.001
					(0.110)	(0.086)	(0.103)	(0.085)
Hr					0.071	0.058	0.099	−0.010
					(0.306)	(0.277)	(0.254)	(0.226)
City fe	YES	YES	YES	YES	YES	YES	YES	YES
Year fe	YES	YES	YES	YES	YES	YES	YES	YES
_cons	0.117***	0.160***	0.084***	0.054***	0.067***	0.126***	0.048***	0.015
	(0.010)	(0.013)	(0.008)	(0.006)	(0.018)	(0.021)	(0.014)	(0.014)
R^2	0.911	0.919	0.916	0.882	0.922	0.931	0.923	0.894
F	3.947**	1.815	3.084*	4.129**	3.319***	4.955***	3.026***	2.564**
N	2,212	2,212	2,212	2,212	2,212	2,212	2,212	2,212

Note:*,**, and***indicate significance at the 0.1, 0.05, and 0.01 levels, respectively, with robust standard errors clustered by city in parentheses. The same applies below.

Furthermore, climate risks exert additional pressure on other cities through the diffusion effects within social systems, further diminishing economic resilience. Extreme weather events typically trigger population migration and resource competition, particularly when residents of disaster-stricken cities move to neighboring regions due to deteriorating living conditions. This migration raises the demand for housing, places additional strain on infrastructure, and increases the cost of living in receiving cities. Such pressure compounds the burden on these cities. Labor force loss may further reduce regional productivity. In addition, financial linkages amplify the social diffusion effect, as financial instability in the market is transmitted through regional banking networks, affecting the financing environment of other cities [35, 36]. Taken together, climate risks have significant impacts on the economic resilience of neighboring cities through both economic and social spillover mechanisms.

Hypothesis 3: Climate risks negatively affect the economic resilience of other cities through spatial spillover effects.

4 Research design

4.1 Sample selection and data sources

Based on the availability of secondary indicators for climate risk and economic resilience, this study selects an initial sample of 168 prefecture-level cities from the period of 2009–2022. The climate risk data is derived from the National Meteorological Centre of China, while data on economic resilience and related control variables are sourced from the "China City Statistical Yearbook." To ensure data integrity and the accuracy of the analysis, the original data were processed as follows: First, linear interpolation was employed to address missing data for some samples; second, samples with significant missing data were excluded; and third, extreme values in the top and bottom 1% and 99% of the main variables were minorized. After these data treatments, a final sample of 158 prefecture-level cities was obtained, resulting in a total of 2,212″city-year" observations.

TABLE 6 Robustness test regression results on the impact of climate risk on urban economic resilience.

	Modification of dependent variables	High-dimensional fixed effects	Lagged explanatory variables	Subsample regression (excluding COVID-19)
	(1)	(2)	(3)	(4)
	CRI	*CRI*	*CRI*	*CRI*
UERI	−1.021**	−0.541**	−0.823***	−0.659*
	(0.453)	(0.270)	(0.267)	(0.334)
Live	−0.007	0.133***	0.147***	0.174***
	(0.011)	(0.033)	(0.035)	(0.038)
Fund	0.001	0.007*	0.008*	0.006
	(0.003)	(0.004)	(0.004)	(0.004)
Den	−0.056	0.270	0.228	0.080
	(0.099)	(0.242)	(0.2425	(0.229)
Export	0.026	0.046	0.052	0.085
	(0.051)	(0.109)	(0.101)	(0.086)
Hr	0.156	0.043	−0.034	−0.302
	(0.236)	(0.307)	(0.314)	(0.384)
_cons	0.362***	1.642	0.080***	0.074***
	(0.017)	(2.275)	(0.019)	(0.018)
City fe	YES	YES	YES	YES
Year fe	YES	YES	YES	YES
Pro-Year fe	NO	YES	NO	NO
R^2	0.895	0.922	0.936	0.928
F	0.904	2.917***	3.683***	3.594***
N	2,212	2,212	2054	1738

4.2 Variable measurement

4.2.1 Explanatory variable: Climate risk index (*CRI*)

The climate risk variable is derived using the methodology of Guo et al. [37], incorporating four sub-indicators: extreme low-temperature days (*LTD.*), extreme high-temperature days (*HTD*), extreme rainfall days (*ERD*), and extreme drought days (*EDD*). The specific processing steps are as follows: first, observations with significant data gaps are removed from the sample; second, historical distribution characteristics for each indicator are calculated from January 1, 1973, to December 31, 1992; third, the frequency and types of extreme weather events at local observation stations from 2009 to 2022 are recorded; finally, the annual extreme weather days are aggregated at the regional level to create a comprehensive Climate Risk Index (*CRI*), as shown in Table 1. To ensure interpretability, the climate risk index is scaled down by a factor of 100. The CRI is computed by the following formula:

$$\mathrm{CRI}_i = \sum_{k=1}^{4} w_k \cdot \mathrm{Sub-Indicator}_k$$

Where: w_k represents the weight assigned to each sub-indicator (LTD., HTD, ERD, EDD), $\mathrm{Sub-Indicator}_k$ is the value of the corresponding sub-indicator. This index aggregates the individual sub-indicators into a single comprehensive measure of climate risk for each region.

TABLE 7 Endogeneity treatment.

	Instrumental variables		Heckman
	(1)	(2)	(3)
	UERI	*CRI*	*CRI*
PM2.5	0.003**		
	(0.001)		
UERI		−0.773**	−0.597**
		(0.302)	(0.254)
IMR			1.735**
			(0.725)
Live	−0.130	0.107	0.742***
	(0.133)	(0.118)	(0.263)
Fund	0.019	0.025	−0.063**
	(0.027)	(0.019)	(0.029)
Den	−2.549**	−1.462	2.130**
	(1.212)	(1.003)	(0.840)
Export	−0.148	−0.148	4.644**
	(0.600)	(0.448)	(1.921)
Hr	0.259	0.721	−2.096**
	(2.395)	(1.802)	(0.9881)
			−1.841**
_cons			(0.800)
City fe	YES	YES	YES
Year fe	YES	YES	YES
Cragg-Donald Wald F statistic	10.610		
Kleibergen-Paap rk LM statistic	6.623[0.011]		
R^2		−2.342	
F		2.6305**	
N	1896	1896	

TABLE 8 Mechanism analysis.

	(1)	(2)
	Psize	*Fin*
UERI	−0.048**	−0.129***
	(0.018)	(0.038)
Live	−0.098	0.145
	(0.096)	(0.139)
Fund	0.0167	−0.007
	(0.022)	(0.029)
Den	−0.859	0.928
	(0.923)	(1.287)
Export	0.021	1.264
	(0.364)	(0.965)
Hr	−1.043	−2.090
	(1.433)	(3.317)
_cons	6.2160***	6.793***
	(0.089)	(0.167)
City fe	YES	YES
Year fe	YES	YES
R^2	0.915	0.752
F	1.562	2.215**
N	2,212	2,212

4.2.2 Dependent variable: Urban economic resilience index (*UERI*)

Urban economic resilience index (*UERI*) is defined as a city's ability to self-recover and dynamically adapt in response to external shocks [38]. This concept encompasses resistance, recovery, and adjustment capabilities. Based on this definition and existing literature, the selection of secondary indicators for *UERI* is grounded in established frameworks for measuring economic resilience and urban development [39]. Specifically, this study measures economic resilience across three dimensions: resistance and recovery capacity (*Rel*), adaptive and adjustment capacity (*Ada*), and innovation and transformation capacity (*Enpu*). Eleven secondary indicators are selected for these dimensions, with indicator weights calculated using the entropy weighting method to construct a comprehensive Urban Economic Resilience Index (*UERI*). The rationale for selecting each indicator is based on its relevance to the city's ability to recover, adapt, and innovate in the face of external shocks, as supported by recent research on urban resilience [38]. Table 2 presents the selected indicators and weights for each dimension.

4.2.3 Mediating variables

4.2.3.1 Urban Population Size (*Psize*)

The size of the urban population serves as a crucial foundation for urban economic development and a key factor in enhancing economic resilience. Following the method of Gao [40], this study

TABLE 9 Spatial spillover effect analysis.

	Adjacency	Geography	Economy
	(1)	(2)	(3)
	CRI	*CRI*	*CRI*
UERI	−0.038***	−0.033***	−0.036***
	(0.011)	(0.012)	(0.013)
Live	0.385***	0.4021***	0.465***
	(0.040)	(0.039)	(0.0402)
Fund	0.067***	0.068***	0.063***
	(0.011)	(0.011)	(0.012)
Den	1.240*	1.405**	2.655***
	(0.705)	(0.712)	(0.810)
Export	0.403	0.751**	0.935***
	(0.327)	(0.326)	(0.343)
Hr	−3.788***	−3.561***	−1.385
	(1.130)	(1.231)	(1.257)
_cons	−0.330***	−0.844***	−0.393***
	(0.115)	(0.229)	(0.131)
Spatial rho	−0.198*** (0.070)	−0.690*** (0.172)	0.067 (0.091)
Variance lgt_theta	−2.196*** (0.169)	−2.219*** (0.169)	−2.452*** (0.160)
sigma2_e	0.001***	0.001***	0.001***
	(0.0001)	(0.0001)	(0.0001)
LR_Direct	−0.039***	−0.034***	−0.036***
	(0.012)	(0.013)	(0.013)
LR_Indirect	0.020 (0.015)	0.019 (0.016)	0.010 (0.019)
LR_Total	−0.019 (0.013)	−0.016 (0.010)	−0.026 (0.0167)
Log-likelihood	715.916	715.269	694.781
R^2	0.642	0.647	0.355

uses the logarithmic form of the year-end total urban population (*LogPsize*) to represent city population size. This transformation effectively reduces the influence of outliers on the analysis and better reflects the general characteristics of urban population scale.

4.2.3.2 Urban Financial Stability (*Fin*)

Urban financial stability measures a city's financial system's capacity to operate under external shocks, especially extreme climate events. This resilience is primarily reflected in the banking system, capital markets, and insurance markets. Drawing on the approach of Xie et al. [41], this study measures financial development level and stability by the ratio of year-end financial institution loan balances to the regional *GDP*.

4.2.4 Control variables

To more comprehensively explore the impact of climate risk on urban economic resilience, this study draws on the control variable design from Xu and Zhang [20], Zhang et al. [42], Zhou et al. [43]. Five control variables are established:

4.2.4.1 Urban Economic Density (*Den*)

Urban economic density reflects the spatial concentration of economic activities. This study represents economic density as the ratio of regional *GDP* to urban land area.

4.2.4.2 Urban Entrepreneurial Activity (*Live*)

Urban entrepreneurial activity indicates the level of dynamism among economic entities in the city. This is measured by the ratio of private and individual business employees to the total urban population.

4.2.4.3 Urban Infrastructure Level (*Fund*)

The level of infrastructure in a city reflects the completeness of its foundational facilities. This study uses *per capita* road area as a measure of urban infrastructure level.

4.2.4.4 Urban Foreign Capital Dependency (*Export*)

Foreign capital dependency indicates the extent to which a city's economy relies on international capital. This is measured by the proportion of actual utilized foreign capital to *GDP*, reflecting the city's dependence on foreign investment.

4.2.4.5 Urban Human Capital Level (*Hr*)

Human capital level is a key indicator of urban workforce quality. This study uses the ratio of enrolled undergraduate and vocational students to the permanent population to represent urban human resource quality.

The data for the indicators are sourced from the China City Statistical Yearbook. The definitions of the variables used in this study are provided in Table 3, and the descriptive statistics for the main variables are shown in Table 4. The descriptive statistics reveal significant heterogeneity among the sample cities with respect to the main variables in the study. Specifically, the mean value of urban economic resilience (*CRI*) is 0.097, with a standard deviation of 0.0947. The maximum and minimum values are 0.9878 and 0.0118, respectively, reflecting notable disparities in the levels of economic resilience across different cities. The mean value of climate risk (*UERI*) is 3.3189, with a standard deviation of 0.2408, indicating that most cities experience climate risk within a narrow range, although some cities are more severely impacted by extreme climate events. The resilience sub-dimensions (*Rel*, *Ada*, *Enpu*) exhibit relatively low mean values, suggesting that cities generally show weaknesses in recovery and resistance, adaptive capacity, and transformational innovation. Control variables, such as urban entrepreneurial vitality (*Live*), infrastructure development level (*Fund*), and economic density (*Den*), display considerable variation, indicating significant

TABLE 10 Regional heterogeneity-1.

	Eastern	Central	Western	Core cities	Non-core cities
	(1)	(2)	(3)	(4)	(5)
	CRI	*CRI*	*CRI*	*CRI*	*CRI*
UERI	−1.281***	−0.159	−0.877	−2.455**	−0.504**
	(0.424)	(0.433)	(0.813)	(1.017)	(0.241)
Live	0.202***	0.060	0.097***	0.101	0.0701***
	(0.064)	(0.040)	(0.035)	(0.079)	(0.018)
Fund	0.021	0.0004	0.009	0.0520*	0.003
	(0.016)	(0.005)	(0.007)	(0.029)	(0.002)
Den	−0.779	0.060	2.055	−3.124	0.008
	(0.987)	(0.073)	(2.081)	(2.988)	(0.198)
Export	0.203	−0.111	0.0800	0.997**	−0.0680
	(0.144)	(0.159)	(0.241)	(0.429)	(0.0790)
Hr	1.410	−0.037	−0.006	−2.944	0.125
	(1.425)	(0.677)	(0.372)	(2.531)	(0.229)
_cons	0.086	0.080***	0.003	0.422*	0.074***
	(0.082)	(0.023)	(0.068)	(0.202)	(0.014)
City fe	YES	YES	YES	YES	YES
Year fe	YES	YES	YES	YES	YES
R^2	0.936	0.878	0.925	0.926	0.918
F	4.744***	0.917	3.081**	2.713**	5.194***
N	728	676	617	234	1787

differences in economic development conditions across cities, which supports the subsequent empirical analysis.

4.3 Econometric model specification

4.3.1 Baseline regression model

To verify the impact of climate risk on urban economic resilience, this study employs a two-way fixed effects model, structured as follows:

$$CRI_{it} = a_0 + a_1 UERI_{it} + a_1 Controls_{it} + \lambda_i + \mu_t + \varepsilon_{it} \quad (1)$$

Where i represents the city, t denotes the year, CRI_{it} indicates the economic resilience of city t in year and $UERI_{it}$ represents the climate risk level for city i in year t. The control variables $Controls_{it}$ include entrepreneurial activity (*Live*), infrastructure development level (*Fund*), economic density (*Den*), foreign capital dependency (*Export*), and human capital level (*Hr*) in city i during year t. λ_i is city-specific fixed effects, which account for unobserved factors that vary across cities but remain constant over time. μ_t is time-specific fixed effects, which capture factors that affect all cities in a given year but vary over time. ε_{it} is the error term, representing unobserved factors that influence the dependent variable but are not included in the model.

4.3.2 Mediation mechanism model

To examine the mechanisms through which climate risk impacts urban economic resilience, this study draws on Jiang and adopts a two-step model specification as follows [44]:

$$\begin{aligned} CRI_{it} &= \gamma_0 + \gamma_1 UERI_{it} + \gamma_i Controls_{it} + \lambda_i + \mu_t + \varepsilon_{it} \\ Mes &= \varphi_0 + \varphi_1 UERI_{it} + \varphi_i Controls_{it} + \lambda_i + \mu_t + \varepsilon_{it} \end{aligned} \quad (2)$$

Here, Med represents the mediation variables, specifically including urban population size (*Psize*) and financial stability (*Fin*).

Other variables remain consistent with Model (1). The two-step process is as follows: The first step involves estimating the baseline model (Equation 1), which measures the direct relationship between climate risk and urban economic resilience. In the second step, the mediation model (Equation 2) is used to assess the indirect impact of climate risk on urban economic resilience through mediation variables, such as population size and financial stability. The coefficient of φ_1, if significantly negative, would indicate that climate risk reduces urban economic resilience by decreasing both urban population size and financial system stability, thus validating Hypothesis 2.

4.3.3 Spatial durbin model (SDM)

To capture the spatial spillover effects of climate risk on urban economic resilience, this study extends Model (1) by introducing a spatial weight matrix and interaction terms among economic resilience, climate risk, and control variables, resulting in the following spatial Durbin model (Equation 3):

$$CRI_{it} = \alpha_0 + \beta_1 W * CRI_{it} + \beta_2 W * UERI_{it} + \alpha_1 UERI_{it} + \beta_1 W * Controls_{it} + \alpha_i Controls_{it} + \lambda_i + \mu_t + \varepsilon_{it} \quad (3)$$

where β_1 represents the spatial autoregressive coefficient. If β_1 is significantly negative, it suggests that climate risk exhibits a spatial spillover effect on urban economic resilience. W denotes the spatial weight matrix, which includes geographic, adjacency, and economic matrices. β_2 and β_1 capture the coefficients of the spatial interaction terms for the explanatory and control variables, respectively. Other variables remain consistent with Model (1) and are not further elaborated here.

5 Empirical analysis

5.1 Baseline regression analysis

In the empirical analysis section, we first conduct a baseline regression analysis to verify whether climate risk suppresses urban economic resilience. Table 5 presents the regression results of baseline model (1), where Columns (1) to (4) show the estimates from the panel fixed-effects model, controlling for both city and time fixed effects but excluding other control variables.

Specifically, Column (1) reports the overall impact of climate risk on the urban economic resilience index, while Columns (2) to (4) explore the effects of climate risk on the three sub-indicators within the resilience capacity system: resistance and recovery capacity (*Rel*), regulation and adaptation capacity (*Ada*), and transformation and innovation capacity (*Enpu*). From these regression results, we can identify the varying impacts of climate risk on each resilience capacity.

Column (2) focuses on the resistance and recovery capacity (*Rel*) sub-indicator, showing a regression coefficient of −0.560, which is not statistically significant. This suggests that climate risk does not significantly affect the resistance and recovery capacity of cities in the sample, indicating that cities' ability to resist and recover from shocks may be less sensitive to climate risks compared to their adaptive or innovative capacities.

In Columns (5) to (8), we further introduce control variables to examine whether the effects of climate risk on the resilience indicators change significantly after accounting for other influencing factors. Based on regression results of Table 5, we preliminarily verify the significant negative impact of climate risk on urban economic resilience. Specifically, Column (1) reports the baseline regression result without including control variables, where the regression coefficient is −0.601 and significant at the 5% level. This indicates that climate risk indeed has a significant suppressive effect on urban economic resilience. After introducing control variables, the result in Column (5) shows a regression coefficient of −0.575, which remains significant. This finding demonstrates robustness and supports Hypothesis 1: climate risk has a direct negative impact on urban economic resilience.

Further, urban economic resilience is decomposed into *Rel, Ada,* and *Enpu* for regression analysis. Column (3) presents the impact of climate risk on urban adaptive and adjustment capacity, with a regression coefficient of −0.455, significant at the 10% level. This suggests that climate risk weakens cities' ability to adapt and adjust. The result reveals that cities face deficiencies in their adaptive and regulatory mechanisms when responding to climate change, which may hinder long-term economic adjustment and transformation goals.

The results in Column (4) indicate that climate risk also has a significant negative impact on urban innovation and transformation capacity. The regression coefficient is −0.373 and significant at the 5% level. This suggests that climate risk not only weakens cities' innovative capabilities but also constrains their potential for future economic transformation.

In Columns (5) to (8), we introduce control variables such as urban entrepreneurial activity, infrastructure development level, economic density, foreign capital dependency, and human capital level. The results show that the negative impact of climate risk on urban resilience remains significant, particularly regarding urban innovation and transformation capacity, where the coefficients remain negative and are significant at the 1% level. Overall, climate risk significantly weakens urban economic resilience, with its impact on innovation and transformation being the most pronounced, further supporting the validity of Hypothesis 1.

5.2 Robustness tests

To ensure a more robust evaluation of the weakening effect of climate risk on urban economic resilience and to further investigate its related impacts, this study conducts robustness tests using the methods of replacing the explained variable, high-dimensional fixed effects, lagged explanatory variables, and subsample regressions. The results are shown in Table 6.

5.2.1 Method of Replacing the Explained Variable

In this test, the explained variable "urban economic resilience" is replaced with "high-quality urban economic development" (measured according to [43]), and the regression results are presented in Column (1) of Table 6. The results indicate that even after substituting urban economic resilience with high-quality urban economic development, the regression coefficient of UERI remains

TABLE 11 Regional heterogeneity - 2.

	Provincial capitals	Non-provincial capitals	Resource cities	Non-resource cities
	(1)	(2)	(3)	(4)
	CRI	*CRI*	*CRI*	*CRI*
UERI	−1.857*	−0.728***	−0.448**	−0.696*
	(0.893)	(0.271)	(0.207)	(0.371)
Live	0.0580	0.140***	0.049**	0.183***
	(0.067)	(0.041)	(0.020)	(0.046)
Fund	0.024*	0.008*	0.005**	0.017*
	(0.013)	(0.005)	(0.002)	(0.010)
Den	1.126	0.098	0.212***	0.655
	(2.916)	(0.185)	(0.031)	(1.325)
Export	−0.532	0.133	0.002	0.095
	(0.443)	(0.090)	(0.117)	(0.134)
Hr	−0.931	−0.272	0.389	−0.299
	(0.974)	(0.287)	(0.240)	(0.486)
_cons	0.147	0.071***	0.053***	0.0420
	(0.182)	(0.018)	(0.009)	(0.071)
City fe	YES	YES	YES	YES
Year fe	YES	YES	YES	YES
R^2	0.891	0.931	0.924	0.927
F	1.473	2.738**	12.426***	2.791**
N	273	1748	784	1,326

significantly negative at the 5% level, demonstrating the robustness of the baseline regression results.

5.2.2 High-Dimensional Fixed Effects Method

To further control for unobservable provincial-level factors that vary over time (such as the impact of macroeconomic changes on urban economic resilience), this study incorporates province-year interaction fixed effects into the baseline regression model. The regression results are shown in Column (2) of Table 6. The results reveal that under the control of high-dimensional fixed effects, the negative impact of climate risk on urban economic resilience remains significant, confirming the robustness of the research conclusions under a higher-dimensional control framework.

5.2.3 Lagged Explanatory Variables Method

Considering the potential continuity and even long-term nature of the impact of climate risk on urban economic resilience, this study lags the climate risk variable by one period to examine the effect of current climate risk on urban economic resilience in the next period. The regression results are presented in Column (3) of Table 6. The findings indicate that after lagging the explanatory variable by one period, the regression coefficient of climate risk remains significantly negative at the 1% level, suggesting that climate risk has a persistent negative effect on urban economic resilience, further confirming its long-term obstructive effect.

5.2.4 Subsample Regression Method:

Given the profound impact of the COVID-19 pandemic on urban economies, particularly in 2020 and thereafter, urban economic resilience during this period may have been affected by the pandemic rather than climate risk. To address this concern, this study excludes samples from 2020 and beyond and reruns the regression analysis. The results are presented in Column (4) of Table 6. The regression results show that even after

removing pandemic-affected samples, the negative impact of climate risk on urban economic resilience remains significant, further supporting the robustness of the research conclusions. In summary, a series of robustness tests further validate the significant negative impact of climate risk on urban economic resilience.

5.3 Endogeneity treatment

The study finds that climate risks significantly reduce urban economic resilience. Conversely, regions with lower economic resilience are more susceptible to severe impacts from climate risks or may consume substantial resources in response to such risks, further hindering resilience improvement. Therefore, a two-way causal endogeneity issue exists. To address the endogeneity problem stemming from this two-way causality, this paper employs the Instrumental Variable (IV) method and the Heckman two-stage method.

First, in the IV method, the PM2.5 index is selected as the instrumental variable. On one hand, the PM2.5 index serves as a crucial indicator for measuring urban climate risk. Over the past two decades, increasing concentrations of PM2.5 have been closely linked to the intensification of environmental stressors and climate-related risks. These elevated pollution levels have contributed to the exacerbation of global climate change, leading to more frequent and severe climate events, such as heatwaves, floods, and storms [45]. As such, the PM2.5 index meets the relevance condition of an instrumental variable by having a strong and established relationship with the urban climate risk. On the other hand, PM2.5 does not directly affect urban economic resilience but may indirectly influence it through pathways such as resource misallocation and population scale, meeting the exclusivity condition. For example, persistent air pollution can drive population migration, alter labor force productivity, or disrupt economic activity by reducing quality of life or escalating health-related expenditures. Hence, the annual median of PM2.5 index is a valid instrumental variable because it satisfies both the relevance and exclusivity conditions, making it an appropriate tool for isolating the causal effect of urban climate risk on economic resilience in the analysis.

The IV regression results are presented in Table 7, Columns (1) and (2). The Cragg-Donald Wald F-statistic is 10.610, ruling out the possibility of weak instruments. The Kleibergen-Paap rk LM statistic is 6.623, with a p-value below 0.05, indicating that the instrumental variable does not suffer from identification issues. Overall, selecting PM2.5 as the instrumental variable is reasonable. The results further reveal that an increase in PM2.5 exacerbates climate risks. In the second-stage regression, the negative impact of climate risks on urban economic resilience remains significant.

Meanwhile, the degree of climate risk may vary across regions due to geographical differences. For example, desert and plateau areas often face more severe climate issues, which have a greater impact on transportation and infrastructure development, thereby exerting a more significant influence on urban economic resilience. This introduces a sample selection bias issue. To mitigate this bias, this paper employs the Heckman two-stage method. Specifically, variables such as urban entrepreneurial vitality (*Live),* infrastructure development level (*Fund*), economic density (*Den*), dependence on foreign capital (*Export*), and human capital level (*Hr*) are selected as the first-stage selection variables to estimate the influencing factors of urban economic resilience in the sample. The inverse Mills ratio (*IMR*), calculated in the first stage, is then included in the second-stage regression analysis. The regression results are shown in Table 7, Column (3). The coefficient of *IMR* is significant at the 5% level, indicating the presence of sample selection bias. After addressing this issue, the regression coefficient of climate risk (*UERI*) remains significantly negative, further supporting the conclusion that climate risks negatively impact urban economic resilience. In conclusion, using both the instrumental variable method and the Heckman two-stage method, this study robustly verifies the significant negative effect of climate risks on urban economic resilience.

5.4 Mechanism analysis

The results above indicate that climate risk significantly suppresses urban economic resilience. What pathways does climate risk use to reduce urban economic resilience? To test the mediating effect of climate risk on urban economic resilience, an empirical analysis of the mediation model was conducted, with results presented in Table 8. Columns (1) and (2) show that climate risk significantly negatively affects population size and financial stability, with statistical significance at the 5% level.

Specifically, climate risk indeed exerts a significant negative impact on urban economic resilience (*CRI*) by reducing urban population size (*Psize*) and disrupting urban financial stability (*Fin*). In detail, the significantly negative influence of climate risk on population size and financial stability further weakens urban economic resilience, thereby validating Hypothesis 2. This finding indicates that climate risk not only directly affects urban economic resilience but also indirectly exacerbates its impact on the urban economic resilience system by influencing population size and financial stability. Therefore, Hypothesis 2 of this paper is supported.

5.5 Analysis of spatial spillover effects

This study adopts the methodology proposed by Elhorst, employing the Lagrange Multiplier (L.M) test, Wald test, SDM model fixed-effects test, Hausman test, and the simplified SDM model test to determine the spatial Durbin model (SDM) with spatial-temporal double fixed effects as the optimal choice for analyzing the spatial spillover effects of climate risk (UERI) on urban economic resilience (CRI). As an extended spatial regression model, the SDM not only estimates the direct effects of climate risk but also captures spatial spillover effects across regions, revealing the transmission pathways of climate risk across different spatial dimensions [46].

From the regression results presented in the Table 9, climate risk demonstrates a significant negative direct impact on urban economic resilience across the dimensions of adjacency, geography, and economic space, with specific coefficients of −0.038, −0.033, and −0.036, all significant at the 1% level. This indicates that climate risk weakens urban economic resilience, regardless of whether the cities are connected through geographic proximity, economic linkages, or spatial adjacency. The spatial correlation coefficient (*rho*) further reveals the spatial spillover effects of climate risk. In

both the adjacency and geographic spatial models, the rho values are negative and significant, indicating a notable negative spatial spillover effect. Notably, in the geographic spatial model, the rho value is −0.690, which suggests that geographically proximate cities are more susceptible to the adverse effects of climate risk.

This finding further demonstrates that climate risk weakens urban economic resilience through negative spillover effects in adjacency, geographic, and economic spaces, with the spillover effect being particularly pronounced among geographically proximate cities. Therefore, Hypothesis 3 is supported.

5.6 Heterogeneous effects of climate risk on urban economic resilience

Due to differences in resource endowments, urban scale, and geographical location, there are significant disparities among cities in terms of both climate risks and economic resilience. This paper further explores the variations in the aforementioned effects. By further exploring these disparities, we can gain a more comprehensive understanding of the vulnerabilities and adaptive capacities of different cities in responding to climate change, thereby providing theoretical support for the formulation of more precise policies and the implementation of corresponding climate adaptation strategies.

5.6.1 Geographical Location

Cities are divided into three groups—eastern, central, and western—based on their geographical location. The analysis results are presented in Columns (1), (2), and (3) of Table 10. The study reveals that the negative impact of climate risks on economic resilience is significantly greater in eastern cities compared to those in central and western regions. A possible reason is that eastern cities have higher levels of economic development and greater dependence on external markets, making them more vulnerable to climate risks. Due to their advanced economic structures and open market environments, eastern regions are more susceptible to disruptions caused by climate events, which substantially weaken their recovery and adaptive capacities.

5.6.2 Urban Hierarchy

Cities are categorized into central and non-central cities based on their importance [47]. Central cities typically refer to those with significant economic status at the regional or national level, such as cities located in economic core areas. These cities possess stronger economic agglomeration effects and policy resources. Non-central cities, on the other hand, have relatively fewer economic activities and resources within their regions. The classification is based on the cities' administrative and functional roles, with central cities playing a more dominant role in regional or national economic development, while non-central cities are more localized in their influence.

The results are shown in Columns (4) and (5) of Table 10. The analysis indicates that the impact coefficient of climate risks on economic resilience in central cities is five times that in non-central cities. Although central cities exhibit stronger recovery capacities, their highly developed economic systems make them more fragile when exposed to climate risks. In contrast, non-central cities, due to their relatively lower levels of economic activity and resource concentration, experience smaller disruptions to their economic resilience systems when facing climate risks.

5.6.3 Administrative Hierarchy of Cities

Cities are classified into provincial capital cities and non-provincial capital cities. The results of the analysis are shown in columns (1) and (2) of Table 11. The negative impact of climate risk is significantly higher for provincial capital cities compared to non-provincial capital cities. Specifically, the climate risk coefficient for provincial capital cities is −1.857, while the coefficient for non-provincial capital cities is −0.728. As regional political, economic, and cultural centers, provincial capital cities experience high resource concentration and dense economic activities, which places them under greater pressure when responding to climate risks. Additionally, these cities bear more responsibilities for policy implementation and infrastructure maintenance, making them more vulnerable to climate shocks. In contrast, non-provincial capital cities have relatively smaller economic scales and fewer policy resources. While they are also affected by climate risks, their resilience systems experience comparatively limited disruption.

5.6.4 Resource Endowment Conditions

According to the National Plan for Sustainable Development of Resource-Based Cities (2013–2020), cities are divided into resource-based and non-resource-based cities. The results are shown in columns (3) and (4) of Table 11. The impact of climate risk differs significantly between resource-based and non-resource-based cities. Non-Resource-based cities may face deficiencies in infrastructure related to climate adaptation and disaster resilience, rendering them more vulnerable to climate risks. For example, excessive exploitation of water resources can lead to insufficient water supply during droughts, thereby affecting both production and daily life. The economic activities of resource-based cities are often directly tied to resource extraction, which entails a higher risk of environmental degradation, making the adverse effects of climate risks particularly pronounced. Furthermore, these cities face challenges such as population aging and shortages of healthcare and educational resources, which exacerbate the impact of climate risks on economic activities and social stability, further weakening their economic resilience.

In summary, the impact of climate risk on urban economic resilience varies significantly based on geographical location, urban hierarchy, administrative level, and resource endowment conditions. These findings suggest that when formulating policies to address climate risks, differentiated strategies should be implemented based on city-specific characteristics. Particular attention should be paid to the vulnerabilities of eastern regions, central cities, provincial capital cities, and resource-based cities to enhance inter-city economic resilience and overall risk resistance capacity.

6 Discussion

Previous literature has largely focused on the internal factors of economic systems that affect economic resilience, with limited attention paid to the influence of external factors on urban

economic resilience, particularly climate risks. This paper, through data collection (Chinese city panel data from 2009–2022) and the application of appropriate empirical analysis methods, tests the hypotheses presented earlier and confirms the reasonableness of the conclusions drawn in the previous literature. In comparison to other related studies, this paper holds significant practical relevance, as with the continuous increase in human activity, especially the expansion of economic activity, environmental sustainability is under profound pressure, and climate issues have become increasingly prominent [24]. Furthermore, the detrimental effect of climate risk on urban economic resilience has emerged as an important issue in contemporary academic discussions. The study's findings suggest that climate risk significantly suppresses urban economic resilience, and this conclusion remains robust across various tests.

On the other hand, this paper adopts more advanced empirical analysis methods, making certain improvements in the construction of the economic resilience evaluation system and the comprehensiveness of the analytical methods, providing conclusions that are timely. The study finds that climate risk weakens urban economic resilience through various mechanisms, consistent with Butsch et al. and Tran and Uzmanoglu [29, 33]. The research highlights that climate risk undermines infrastructure, leads to population loss, and causes financial system fragility, thereby impacting urban economic resilience. Additionally, climate risk has a significant negative impact on urban economic resilience across neighboring, geographical, and economic dimensions, suggesting that its effects are not confined to the affected cities but spread to surrounding areas through economic networks and supply chains. This finding is also consistent with Nyarko's research on supply chain disruptions and resource competition [34]. Given the different factor endowments across regions, the impact of climate risk on urban economic resilience also varies significantly.

7 Conclusions and policy implications

7.1 Conclusions

In the context of intensifying global climate risks and growing attention on urban economic resilience, this paper examines the impact of climate risk on the economic resilience of Chinese cities and its underlying mechanisms. Using panel data from Chinese cities between 2009 and 2022, the paper constructs an index system for urban economic resilience and assesses the level of climate risk in each city. Through empirical analysis, the study explores both the direct impact of climate risk on urban economic resilience and its indirect effects mediated by factors such as urban population size and financial stability. In addition, the spatial spillover effects of climate risk are analyzed, revealing the varying impacts across different regions and city types. The results show that climate risk significantly weakens urban economic resilience, a conclusion that remains robust after conducting various robustness checks. Furthermore, climate risk exacerbates this impact by influencing urban population size and financial stability. The effect of climate risk varies by region and city type: compared to the central and western regions, non-central cities, and resource-based cities, the eastern region, central cities, and non-resource-based cities are less affected by climate risk. Finally, the study identifies clear spatial spillover effects, where climate risk impacts do not remain confined to individual cities but spread to surrounding cities or regions through economic connections, supply chains, and market fluctuations.

7.2 Limitations and future research directions

Although this study strives to ensure the reliability of the results, there are still some limitations. First, while multiple control variables, regression models, and spatial Durbin models have been used to mitigate the impact of endogeneity, it is not possible to completely rule out the potential correlation between explanatory variables and error terms. This endogeneity may arise from bidirectional causality, omitted variables, or measurement errors, such as the possibility that urban economic resilience may, in turn, influence its capacity to withstand climate risks. Furthermore, the exclusion of key variables such as government policy, economic crises, natural disasters and public health emergencies may lead to bias in the estimates of climate risk effects, impacting the accuracy of the conclusions. Future research could further alleviate these issues by incorporating instrumental variables or quasi-natural experiments and including the integration of longer time series and social events in the analysis. Second, given that this study primarily focuses on Chinese cities, China's unique political system may influence the relationship between climate risk and economic resilience. The differences in regional policies may affect resource allocation and implementation strength, thereby impacting the enhancement of urban economic resilience. For example, the eastern region emphasizes high-quality development and ecological protection, while the central and western regions focus on economic growth and infrastructure development. Third, the public opinion to climate risks is an important driving force for governments to take measures to address climate change and economic resilience. In future research, we will further consider the impact of public opinion. Cross-national comparisons, studies on behavior and attitudes, as well as research on social movements and policy change, will help us gain a more comprehensive understanding of and effectively utilize public opinion as an important force to better address the global climate change challenge.

7.3 Policy implications

This study contributes to the theoretical understanding of urban economic resilience by emphasizing the complex and multifaceted impacts of climate risks on cities. Based on the above conclusions and discussions, this paper provides several insights for policymakers:

7.3.1 Increase investment in climate-adaptive infrastructure to enhance the city's ability to respond to climate risks

To address the challenges posed by climate change, the government should prioritize investments in infrastructure that

can mitigate extreme weather events, such as flood control, drought prevention, and drainage systems. In addition, projects like smart transportation and green buildings should be promoted to improve the sustainability of urban infrastructure. For non-central cities and resource-based cities, it is crucial to prioritize the construction of disaster-resistant infrastructure to ensure the cities can recover swiftly when climate-related events occur. Furthermore, the use of clean and renewable energy should be encouraged to reduce dependence on fossil fuels and improve the long-term sustainability of cities.

7.3.2 Promote the diversification of urban economic structures to enhance urban economic resilience

Climate risks have a more pronounced impact on cities with a single economic structure. Policies should encourage cities in the central and western regions, as well as resource-based cities, to adjust their industrial structures and drive technological innovation. The government should support the development of green industries, low-carbon technologies, and renewable energy through tax incentives, financial subsidies, and other measures. Furthermore, innovation by high-tech companies should be encouraged to steer the industrial structure towards low-carbon and environmentally friendly directions. Additionally, it is essential to strengthen urban development in the fields of digital economy and green buildings to improve the overall economic resilience of cities, thereby maintaining competitiveness in the context of climate change.

7.3.3 Establish cross-regional cooperation mechanisms to enhance the synergistic effect of climate risk management

Due to the spatial spillover effect of climate risks, cities should strengthen regional coordination and establish cross-regional climate risk management mechanisms. Developed regions in the east can leverage their economic and technological advantages to assist cities in the central and western regions, as well as non-central cities, in enhancing their climate adaptation capabilities, thereby forming a collaborative regional network for responding to climate change. Additionally, cross-regional climate emergency response and resource-sharing platforms should be established to facilitate the sharing of climate risk information and the collaborative execution of emergency plans.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: China City Statistical Yearbook: https://www.stats.gov.cn/sj/ndsj/. The National Meteorological Centre of China https://k.data.cma.cn/mekb/r=site/indexdata.

Author contributions

DW: Writing – original draft, Writing – review and editing, Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software. QA: Funding acquisition, Project administration, Resources, Supervision, Validation, Visualization, Writing – review and editing. LZ: Project administration, Resources, Validation, Visualization, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

This article was submitted as a preprint to Preprints.org as Wu et al. [49].

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. Used by non-native English writers when doing enhancement writing.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Zhang J, Yang Z, He B. Empowerment of digital Technology for the Resilience of the logistics industry: mechanisms and paths. *Systems* (2024) 12(8):278. doi:10.3390/systems12080278

2. Ding YJ, Li CY, Xiaoming W, Wang Y. An overview of climate change impacts on society in China. *Adv Clim Change Res* (2021) 12(2):123–34. doi:10.1016/j.accre.2021.03.002

3. Abid N, Ahmad F, Aftab J, Razzaq A. A blessing or a burden? Assessing the impact of climate change mitigation efforts in Europe using quantile regression models. *Energy Policy* (2023) 178:113589. doi:10.1016/j.enpol.2023.113589

4. Kashi SMH, Farrokhzadeh S, Baharvandi S, Zolfani SH. Effects of extreme weather events and climate change on cities' livability. *Cities* (2024) 151:105114. doi:10.1016/j.cities.2023.105114

5. Zhang J, Yang Z, He B. Does digital infrastructure improve urban economic resilience? Evidence from the Yangtze River Economic Belt in China. *Sustainability* (2023) 15(19):14289. doi:10.3390/su151914289

6. He B, Xu D, Nan G, Zhang X, Yu X. Does the cross-border e-commerce comprehensive pilot zones policy affect the urban–rural income gap in China? *Am J Econ Sociol* (2024) 83(4):773–92. doi:10.1111/ajes.12593

7. Briguglio L, Cordina G, Farrugia N, Vella S. Economic vulnerability and resilience: concepts and measurements. *Oxford Development Stud* (2009) 37(3):229–47. doi:10.1080/13600810903089893

8. Liu J, Huang X, Chen J. High-speed rail and high-quality development of urban economy: an empirical study based on the data of prefecture-level cities. *Contemp Finance and Econ* (2021) (1) 14–26. doi:10.13676/j.cnki.cn36-1030/f.2021.01.003

9. Wang Z, He B, Fujita T. Does environmental regulation affect inward foreign direct investment? Evidence from China. *Appl Econ Lett* (2024) 1–7. doi:10.1080/13504851.2024.2352164

10. Ma J, Caldecott B, Volz U. Case studies of environmental risk analysis methodologies. *Environ Risk Anal Rev* (2020) (9). doi:10.19409/j.cnki.thf-review.2020.09.003

11. Campiglio E, Daumas L, Monnin P, von Jagow A. Climate-related risks in financial assets. *J Econ Surv* (2023) 37(3):950–92. doi:10.1111/joes.12525

12. Carleton TA, Hsiang SM. Social and economic impacts of climate. *Science* (2016) 353(6304):aad9837. doi:10.1126/science.aad9837

13. Hu Z, Borjigin S. The amplifying role of geopolitical risks, economic policy uncertainty, and climate risks on energy-stock market volatility spillover across economic cycles. *The North Am J Econ Finance* (2024) 71:102114. doi:10.1016/j.najef.2024.102114

14. Lin B, Wu N. Climate risk disclosure and stock price crash risk: the case of China. *Int Rev Econ and Finance* (2023) 83(C):21–34. doi:10.1016/j.iref.2022.08.007

15. Zhang D, Jiang D, He B. Empowering agricultural economic resilience with smart supply chain: theoretical mechanism and action path. *Sustainability* (2025) 17(7):2930. doi:10.3390/su17072930

16. Anh DLT, Anh NT, Chandio AA. Climate change and its impacts on Vietnam agriculture: a macroeconomic perspective. *Ecol Inform* (2023) 73:101960. doi:10.1016/j.ecoinf.2022.101960

17. Löscher A, Kaltenbrunner A. Climate change and macroeconomic policy space in developing and emerging economies. *J Post Keynesian Econ* (2022) 46(1):113–41. doi:10.1080/01603477.2022.2084630

18. Sun J, Sun X. Progress of regional economic resilience research and exploration of its application in China. *Econ Geogr* (2017) 37(10):1–9. doi:10.15957/j.cnki.jjdl.2017.10.001

19. Liu L, Meng Y, Wu D, Ran Q, Cao J, Liu Z. Impact of haze pollution and human capital on economic resilience: evidence from prefecture-level cities in China. *Environ Development Sustainability* (2023) 25:13429–49. doi:10.1007/s10668-022-02625-8

20. Xu Y, Zhang L. The economic resilience and origin of Chinese cities: the perspective of industrial structure diversification. *Finance and Trade Economy* (2019) 40(7):110–26. doi:10.19795/j.cnki.cn11-1166/f.2019.07.008

21. Zhao R, Li X, Wang Y, Xu Z, Xiong M, Jia Q, et al. Assessing resilience of sustainability to climate change in China's cities. *Sci The Total Environ* (2023) 898:165568. doi:10.1016/j.scitotenv.2023.165568

22. Wu Z, Sun S, Zhang S, Wang W. Assessment of climate resilience of cities in the Yellow River Basin and countermeasures to improve it. *Land Resour Sci Technology Management* (2024) 41(2):62–74. doi:10.3969/j.issn.1009-4210.2024.02.006

23. He B, Yu X, Ho SE, Zhang X, Xu D. Does China's two-way FDI coordination improve its green total factor productivity? *Polish J Environ Stud* (2024) 33(1):173–83. doi:10.15244/pjoes/171681

24. Zhang L, Xu M, Chen H, Li Y, Chen S. Globalization, green economy and environmental challenges: state-of-the-art review for practical implications. *Front Environ Sci* (2022) 10:870271. doi:10.3389/fenvs.2022.870271

25. Salimi M, Al-Ghamdi SG. Climate change impacts on critical urban infrastructure and urban resiliency strategies for the Middle East. *Sustainable Cities Soc* (2020) 54:101948. doi:10.1016/j.scs.2019.101948

26. Parrado R, Bosello F, Delpiazzo E, Hinkel J, Lincke D, Brown S. Fiscal effects and the potential implications on economic growth of sea-level rise impacts and coastal zone protection. *Climatic Change* (2020) 160(2):283–302. doi:10.1007/s10584-020-02664-y

27. Perera ATD, Hong T. Vulnerability and resilience of urban energy ecosystems to extreme climate events: a systematic review and perspectives. *Renew Sustainable Energy Rev* (2023) 173:113038. doi:10.1016/j.rser.2022.113038

28. Wang X, Li M. Determinants of regional economic resilience to economic crisis: evidence from Chinese economies. *Sustainability* (2022) 14(2):809. doi:10.3390/su14020809

29. Butsch C, Beckers LM, Nilson E, Frassl M, Brennholt N, Kwiatkowski R, et al. Health impacts of extreme weather events – cascading risks in a changing climate. *J Health Monit* (2023) 8(Suppl. 4):33–56. doi:10.25646/11652

30. Xie M, Feng Z, Li C. How does population shrinkage affect economic resilience? A case study of resource-based cities in Northeast China. *Sustainability* (2022) 14(6):3650. doi:10.3390/su14063650

31. Abe M, Ye L Building resilient supply chains against natural disasters: the cases of Japan and Thailand. *Glob Business Rev* (2013) 14(4):567–86. doi:10.1177/0972150913501606

32. Wang Z, Wei W. Regional economic resilience in China: measurement and determinants. *Reg Stud* (2021) 55(7):1228–39. doi:10.1080/00343404.2021.1872779

33. Tran N, Uzmanoglu C. Climate risk and credit ratings. *J Financial Res* (2024) 47:947–80. doi:10.1111/jfir.12388

34. Nyarko IK. The impact of natural disasters on international trade: the potential effects of April 2024 Dubai floods on Ghana's economy. *Business, Management Econ Res Prog* (2024) 4:64–86. doi:10.9734/bpi/bmerp/v4/1540

35. West CD, Stokeld E, Campiglio E, Croft S, Detges A, Duranovic A, et al. Europe's cross-border trade, human security and financial connections: a climate risk perspective. *Clim Risk Management* (2021) 34:100382. doi:10.1016/j.crm.2021.100382

36. Li N. Measuring the combining effects of financial stability and climate risk for green economic recovery. *Econ. Change Restruct.* (2023) 56(2):1225–1241. doi:10.1007/s10644-022-09466-5

37. Guo K, Ji Q, Zhang D. A dataset to measure global climate physical risk. *Data in Brief* (2024) 54:110502. doi:10.1016/j.dib.2024.110502

38. Martin R, Sunley P, Gardiner B, Tyler P. How regions react to recessions: resilience and the role of economic structure. *Reg Stud* (2016) 50(4):561–85. doi:10.1080/00343404.2015.1136410

39. Béné C, Al-Hassan RM, Amarasinghe O, Fong P, Ocran J, Onumah E, et al. Is resilience socially constructed? Empirical evidence from Fiji, Ghana, Sri Lanka, and Vietnam. *Glob Environ Change* (2016) 38:153–70. doi:10.1016/j.gloenvcha.2016.03.005

40. Gao H. Urban population size and labor income. *The J World Economy* (2014)(10) 145–64. doi:10.19985/j.cnki.cassjwe.2014.10.008

41. Xie J, Zheng Y, Dong Q. The digital and intelligent construction of supply chains enables the new quality productive forces of manufacturing enterprises: a quasi-natural experiment based on the construction of supply-chain innovation and application pilot cities. *J Shanghai Univ Finance Econ* (2024) 26(5):15–29. doi:10.16538/j.cnki.jsufe.2024.05.002

42. Zhang Y, Deng S, Zhang Y. Research on urban economic resilience: theoretical progress and future direction. *J Management* (2022)(2) 54–67. doi:10.19808/j.cnki.41-1408/f.2022.0015

43. Zhou Q, Zhu M, Qiao Y, Zhang X, Chen J. Achieving resilience through smart cities? Evidence from China. *Habitat Int* (2021) 111:102348. doi:10.1016/j.habitatint.2021.102348

44. Jiang T. Mediating effects and moderating effects in causal inference. *China Ind Econ* (2022) 5(100):r120. doi:10.19581/j.cnki.ciejournal.2022.05.005

45. Lim CH, Ryu J, Choi Y, Jeon SW, Lee WK. Understanding global PM2.5 concentrations and their drivers in recent decades (1998–2016). *Environ Int* (2020) 144:106011. doi:10.1016/j.envint.2020.106011

46. Elhorst JP. Matlab software for spatial panels. *Int Reg Sci Rev* (2014) 37(3):389–405. doi:10.1177/0160017612452429

47. Fang C, Yu D. Urban agglomeration: an evolving concept of an emerging phenomenon. *Landscape Urban Plann* (2017) 162:126–36. doi:10.1016/j.landurbplan.2017.02.014

48. Gong M, Huang X. Progress in industry chain resilience research. *SHS Web of Conferences* (2023) 163:02013. doi:10.1051/shsconf/202316302013

49. Wu D., Zheng L., An Q. (2025). Does Climate Risk Undermine Urban Economic Resilience?—Evidence from Chinese Cities. *Preprints*. 2025041642. doi:10.20944/preprints202504.1642.v1

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 27 June 2025
DOI 10.3389/fphy.2025.1609866

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Yasuko Kawahata,
Rikkyo University, Japan
Jiusheng Chen,
Chongqing Normal University, China

*CORRESPONDENCE
Zhong Wang,
wangzhong2012@cdut.edu.cn
Yi Wang,
Wyde1028798@163.com

RECEIVED 11 April 2025
ACCEPTED 19 May 2025
PUBLISHED 27 June 2025

CITATION
Duan Y, Jing X, Xu H, Wang Z, Zhang X and Wang Y (2025) Complex impacts between economic policy and listed oil companies: a dynamics-based approach.
Front. Phys. 13:1609866.
doi: 10.3389/fphy.2025.1609866

COPYRIGHT
© 2025 Duan, Jing, Xu, Wang, Zhang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Complex impacts between economic policy and listed oil companies: a dynamics-based approach

Yanzhi Duan[1], Xingsheng Jing[2], Haitang Xu[2], Zhong Wang[3]*, Xin Zhang[1] and Yi Wang[3]*

[1]Research Institute of Natural Gas Economics, PetroChina Southwest Oil & Gasfield Company, Chengdu, China, [2]PetroChina Southwest Oil & Gasfield Company, Chengdu, China, [3]College of Management Sciences, Chengdu University of Technology, Chengdu, China

The causal impacts between economic policy and listed oil companies have been discussed previously, reporting either positive or negative impacts. However, various channels bridge the impacts flowing between them, resulting in multiple causal impacts rather than only a single causal impact. This work employs a dynamics-based causality method, namely, pattern causality, and aims to reveal multiple causal interactions between economic policy and listed oil companies in China. Using the economics policy uncertainty index and the stock prices of nine popular listed oil companies in China from 2006 to 2025, several interesting findings are discovered. 1) Three kinds of causalities exist between economic policy and listed oil companies: positive, negative, and dark causality. Dark causality suggests more complex interactions, which may drive two variables to evolve with fluctuations; 2) rather than only one kind of causality, three heterogeneous causalities exist simultaneously with different magnitudes; 3) the causal impacts between the economic policy and listed oil companies are asymmetric; 4) by comparing the causal magnitudes, dark causality is shown to play a relatively dominant role between economic policy and listed oil companies. These new findings suggest complex interactions between economic policy and listed oil companies, requiring a reassessment of the underlying risks for market participants. Our approaches also offer an alternative perspective on understanding the complexity of social systems.

KEYWORDS

economic policy, oil listed company, multiple causality, dynamics theory, complex system analysis

1 Introduction

The global energy landscape is undergoing a profound transformation, shaped by overlapping forces such as climate change mitigation efforts [1, 2], geopolitical realignments [3, 4], and technological innovation [5]. Central to this transformation is the interplay between economic policy uncertainty (EPU) and the strategic behavior of listed oil companies [6, 7]. As governments grapple with decarbonization targets, energy security concerns, and macroeconomic stabilization, the resultant policy volatility creates both risks and opportunities for oil firms [8]. Conversely, the operational decisions of these firms—ranging from capital expenditure allocations to lobbying activities—significantly

influence policy trajectories [9, 10]. This paper explores the bidirectional dynamical impacts between EPU and listed oil companies in China, addressing a critical gap in the literature by analyzing the dynamic feedback mechanisms that characterize their interactions.

Economic policy uncertainty, defined as the ambiguity surrounding future policy actions [11], has surged in recent decades due to factors such as trade wars, regulatory reforms, and the transition to net-zero economies [12, 13]. For the oil sector, which is capital-intensive, politically sensitive, and exposed to long-term demand risks, EPU exacerbates operational challenges [14]. For example, the EU's evolving carbon pricing framework and the US's vacillating stance on shale oil subsidies have created a climate of uncertainty that disrupts investment planning (BP, 2022). Conversely, oil company responses—such as divesting from fossil fuels or expanding into renewable energy—can alter policymakers' perceptions of industry viability, thereby shaping subsequent regulatory choices [13].

Existing studies on EPU and oil firm behavior primarily focus on directional impacts [15, 16]. For instance, a large body of research documents how EPU reduces corporate investment [17], increases financing costs [18], and distorts innovation trajectories [19]. In the energy sector, scholars have examined how policy uncertainty affects renewable energy adoption (Aklin & Urpelainen, 2012) and oil price volatility [20, 21]. However, these studies often treat policy as an exogenous shock, neglecting the role of firms in shaping policy outcomes. The findings are mixed regarding the impacts between EPU and oil firms. Some scholars report negative causal impacts between EPU and firm behavior [22]. For example, some scholars suggest that policy uncertainty may increase uncertainty as it leads to increased costs and, therefore, negatively affects investment behavior. Bloom [23] has reported that increased uncertainty raises the "real option value to wait," which forces companies to delay their investment until the uncertainty is reduced. In contrast, some literature provides evidence of positive causal impacts. Belderbos et al. [24] reported that firms in across-country investments improve the value of growth options, and this often occurs when high market uncertainty exists in host countries. In summary, the impacts between EPU and listed oil companies are debatable and mixed concerning the directional and bidirectional as well as positive and negative impacts, requiring new techniques to clarify the complex connections between EPU and oil firms.

The interplays between EPU and listed oil companies may occur from various channels, as shown in Figure 1 [25–27]. 1) For investment decision-making. Policy uncertainty increases the unpredictability of a company's future cash flows, leading oil companies to postpone or reduce long-term capital expenditures such as exploration and development and refining capacity expansion [28]. 2) Financing cost channel. The rise in market risk premiums caused by policy fluctuations is transmitted through debt financing costs and equity financing costs [22]. 3) Expectation management [29]. The ambiguity of policy signals leads to deviations in management's forecasts of future demand. OPEC + member countries adjusted their monthly production quotas repeatedly due to the repeated release policy of the US Strategic Petroleum Reserve. 4) Supply chain reconfiguration. Changes in trade policies (such as the US sanctions on Venezuelan crude oil) force oil companies to adjust their procurement networks, increasing logistics costs



FIGURE 1
Possible impact patterns and channels between EPU and listed oil companies.

[30]. 5) Tax policy. Changes in policies such as windfall profit taxes directly affect corporate profits (for example, the UK's Energy Profits Levy in 2022 led to a 12% decline in BP's quarterly net profit.) 6) Technological innovation. The uncertainty of the carbon pricing mechanism affects the R&D investment of oil companies in carbon capture, utilization, and storage (CCUS) technologies [31]. 7) Market access. The tightening of environmental protection policies (such as the EU's plan to ban the sale of fuel-powered vehicles by 2030) forces oil companies to adjust their business structures. 8) Consumer expectation [32]. The expected fluctuations in energy prices triggered by policy uncertainty affect terminal demand [33]. 9) Political risk channel [21, 34, 35]. Geopolitical conflicts (such as the Russia–Ukraine war) combined with the uncertainty of energy policies led to the revaluation of the assets of oil companies. These channels form a dynamic feedback mechanism. For example, oil companies influence policy-making through political donations, thus forming a two-way interaction between policies and corporate behavior. It is obvious that some channels may pass positive impacts between economic policy and listed oil companies, while negative impacts flow through other channels. These facts suggest not simply a single causal impact between economic policy and listed oil companies. Moreover, the fluctuation trends between variables are often neither the same nor opposite. In addition to the positive and negative impact patterns, more complex patterns may exist between EPU and listed oil companies.

Some scholars aim to detect the interplay and volatility spillover between economic policy and listed oil companies [36, 37]. Another popular approach is causality analysis [38–40]. The Granger causality test is one of the popular causality analysis methods, and it is widely used in economics and finance [41]. Due to the nonlinearity in many real-world systems, the nonlinear Granger causality test was developed to measure nonlinear causal impacts [42]. However, traditional Granger causality methods are based on statistic regression and are reliable for stable series, but they show limitations for unstable series. Moreover, they also show limitations when we must quantify the strength of causal impacts. Transfer entropy is a nonparametric tool that investigates causality by measuring the information flowing between observations. As certified by previous works, transfer entropy is equal to the Granger causality test when the sample datasets are Gaussian distributed [43]. Unlike previous statistical approaches, cross-convergent mapping (CCM) is developed from the dynamics embedded in the state space [44]. It reports that the time series variables are causally linked if they are generated from the same dynamical system and they share a common attractor. The CCM

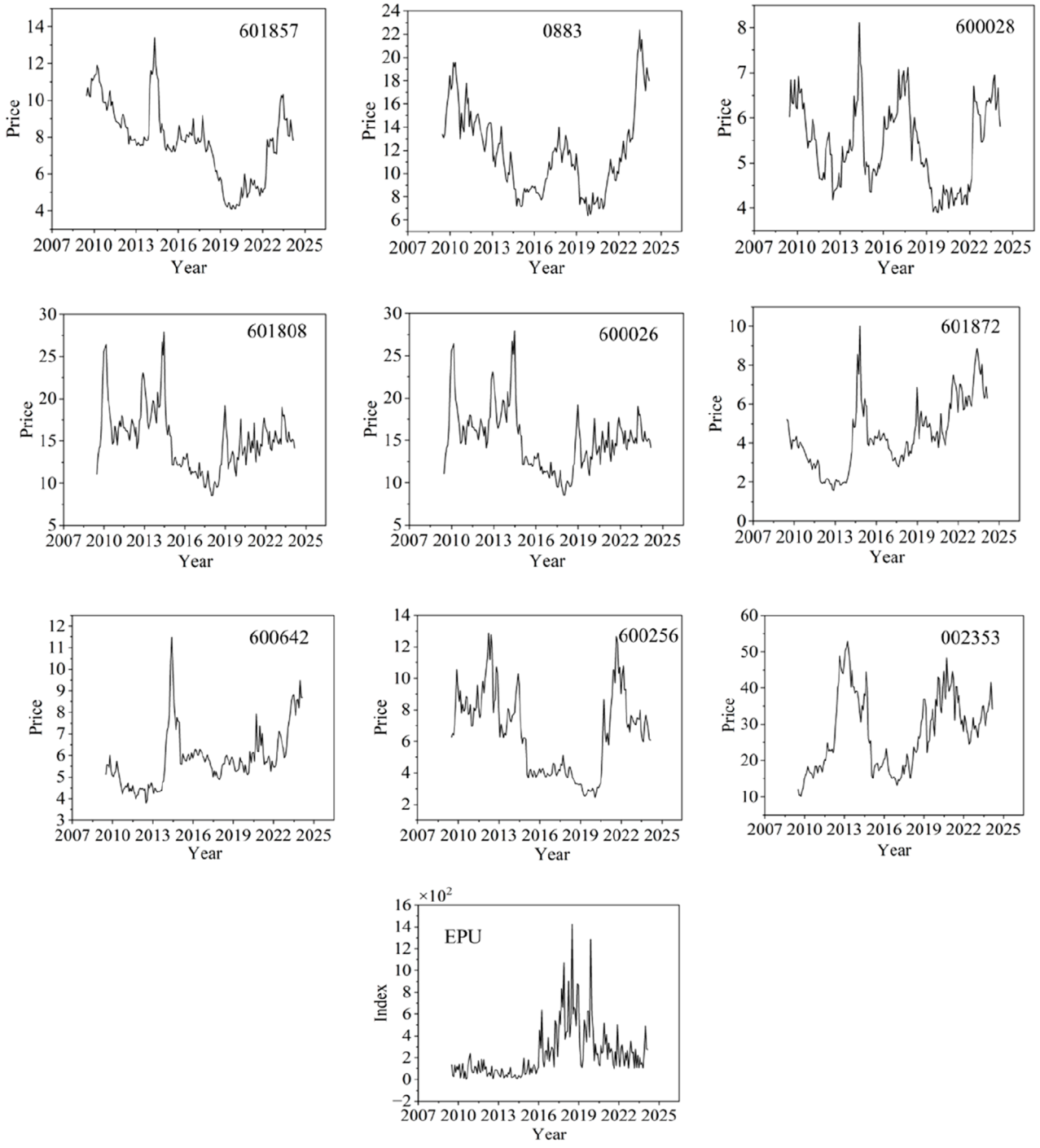


FIGURE 2
The selected sample data in this work.

method has a satisfactory ability to detect causality in complex nonlinear systems, which solves the limitation of the Granger causality test.

Although several causal detection techniques are widely used in financial systems, like Granger causality, transfer entropy, and CCM, these methods can only distinguish a single causal impact. For example, only positive or negative causal impacts can be obtained between two observations *via* traditional methods. This is very limited for practical situations. If there is only a single causal impact (e.g., positive impact) from variable X to variable Y, then the increase (decrease) of X will drive the increase (decrease) of variable Y. In other words, they are more likely to evolve in similar directions. However, these rules are often rare for many real-world observations; their fluctuation associations

TABLE 1 Statistical analysis of sample datasets.

Order	Code	Mean	STD	Maximum	Minimum	Skewness	Kurtosis
1	601857	7.81	20.3	13.39	4.07	0.05	−0.42
2	0883	11.86	3.63	22.4	6.37	0.71	−0.1
3	600028	5.4	0.91	8.12	3.9	0.22	−0.84
4	601808	15.21	3.66	27.94	8.54	0.92	1.36
5	600026	7.96	3.47	18.09	3.33	1.04	0.13
6	601872	4.42	1.72	10	1.59	0.57	−0.01
7	600642	5.86	1.32	11.48	3.8	1.32	2.29
8	600256	6.48	2.58	12.88	2.44	0.3	−0.78
9	002353	27.93	10.49	52.94	10.08	0.32	−0.95
10	EPU	234.9932	237.75	1425.2	0.33	2.08	5.67

Note: The stock prices of China Petrochemical Corporation (0083) are collected from Hong Kong Exchanges and Clearing Limited. The stock prices of Jereh Group (002353) are collected from the Shenzhen Stock Exchange. Other stock prices are collected from the Shanghai Stock Exchange.

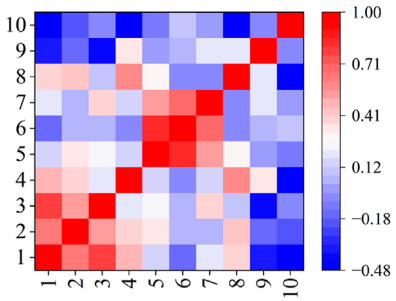


FIGURE 3
The Pearson correlation analyses of selected samples.

are complex, neither the same nor opposite directions. The complex causality between them can be partially explained by the aforementioned various channels. Some channels propagate positive impacts, while others flow through negative causal impacts. Thus, we speculate that there may be more than one kind of causal impact between economic policy and listed oil companies and possibly multiple impacts. Thus, more reliable methods are needed to reveal the complex interactions in real-world systems. Stavroglou et al. [45] proposed the pattern causality algorithm (PC) based on the state space reconstruction theory, which can not only detect causality, but it can also quantify causal strength. More importantly, PC can distinguish both positive and negative causal impacts as well as a more complex causal interaction, namely, dark causality.

Thus, this work employs pattern causality methods to examine multiple causality impacts between economic policy and listed oil companies. The study makes three contributions. First, it utilizes a new theoretical framework that captures the complex dynamical causalities between EPU and listed oil firms in China. Second, in addition to the previously reported positive and negative causality, it reveals more complex causal interactions (dark causality) between EPU and the stock prices of listed companies, both positive and negative. Third, it analyzes the dominant causal interactions between economic policy and listed oil companies by quantifying the magnitude of multiple causalities.

The remainder of this paper is organized as follows: Section 2 describes the methods. Section 3 depicts the sample data. In Section 4, we investigate the complex causal impacts between EPU and the stock prices of listed oil companies in China. Section 5 summarizes our main findings.

2 Methods

According to dynamical systems theory, in a dynamic system, if we can estimate the state of component Y based on the dynamics of component X, we can infer causality from variable Y to variable X [46]; the more accurate the estimation can be, and the stronger the causality is. On the other hand, inaccurate estimation indicates variable Y has no dynamics information left in variable X. We speculate that there is no causal interaction from Y to X. According to [44], it is possible to test estimations by the mappings between reconstructed dynamical structures from time series variables. In particular, we can measure the causality by mapping among neighbor points in the reconstructed dynamical structure. The dynamical structure is reconstructed based on the delay embedding theory. Thus, we can test the dynamical causal interactions between variables based on the accuracy of mutual estimations. As we know, the causal interactions between variables can be positive or negative, even in more complex ways. In addition, the strengths of impacts are different for different samples, thus

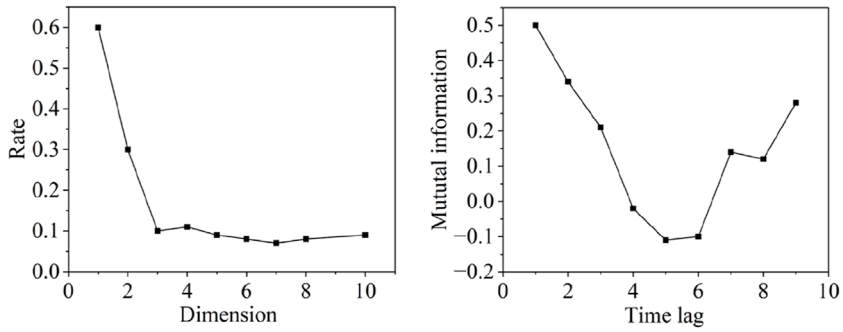


FIGURE 4
The selection of parameters for the reconstructions.

requiring not only causality inference but also the causal strength and causal type.

A dynamics-based causality method is proposed to solve the above issues, namely, pattern causality, which examines causality from X to Y by quantifying how consistently average patterns in the state space of X correspond to average patterns in the state space of Y in contemporaneous neighborhoods. According to pattern causality, there may exist three heterogeneous causal impacts: positive, negative, and dark causalities. A positive causality (i.e., "mutualism") drives two variables to evolve in the same direction, while opposite forces exist for negative causality (i.e., "competition"). Dark causality suggests more complex interactions that are neither "mutualism" nor "competition," and it forces two variables to evolve with fluctuations.

Specifically, the main processes of pattern causality are given by four fundamental steps [40, 45]. First, we must make a state space reconstruction from time series [47]. Based on Takens' theory, we can reconstruct a multidimensional state space from time series using two faithful parameters, that is, embedding dimension E and lag τ [48]. Specifically, we reconstruct E-dimensional state space A_X from time series $X(t)$, see Equation 1.

$$A_X = \begin{bmatrix} X(1) & X(1+\tau) & \cdots & X(1+(E-1)\tau) \\ X(2) & X(2+\tau) & \cdots & X(2+(E-1)\tau) \\ \vdots & \vdots & \ddots & \vdots \\ X(h) & X(h+\tau) & \cdots & X(l) \end{bmatrix}, \quad (1)$$

where l is the length and $h = l-(E-1)\tau$. The row vectors $(x(t))$ on A_X denote the states evolving with time.

Once we get the reconstructed state space for both variables X and Y, we must then determine the average pattern of each point in the state space. The average pattern shows the fluctuation trend in state space. The average pattern of $x(t)$ is estimated by its nearest neighbors. Specifically, we must find the weighted combination from its E+1 nearest points based on the Euclidean distance, see Equation 2:

$$w_i^x = \frac{e^{-d(x(t),x(i))}}{\sum_i e^{-d(x(t),x(i))}}, \quad (2)$$

where d denotes the Euclidean distance, and $x(i)$ is the nearest neighbor of a given $x(t)$.

The average pattern $P_{x(t)}$ can be determined by the coarse process; see Equations 3–5.

$$h_i^x = \left(\frac{X(t_{x(i)}+\tau) - X(t_{x(i)})}{X(t_{x(i)})}, \cdots, \frac{X(t_{x(i)}+(E-1)\tau) - X(t_{x(i)}+(E-2)\tau)}{X(t_{x(i)}+(E-2)\tau)} \right), \quad (3)$$

$$H_{x(t)} = \sum_{i=1}^{E+1} w_i^x h_i^x, \quad (4)$$

$$P_{x(t)} = Pattern\left(H_{x(t)}\right). \quad (5)$$

Nine patterns exist in a 3-dimensional state space: $(\nearrow,\nearrow)$, $(\nearrow,\rightarrow)$, $(\nearrow,\searrow)$, $(\rightarrow,\nearrow)$, $(\rightarrow,\rightarrow)$, $(\rightarrow,\searrow)$, $(\searrow,\nearrow)$, $(\searrow,\rightarrow)(\searrow,\searrow)$.

Third, we distinguish three kinds of causality based on the average pattern in the state space A_X and A_Y. Positive causality is inferred when A_Y has a similar average pattern on A_X. Negative causality is determined when opposite patterns are coupled; for example, a $(\nearrow,\nearrow)$ pattern causes the $(\searrow,\searrow))$ pattern. Moreover, a dark causality is inferred when neither similar nor opposite patterns (e.g., pattern $(\nearrow,\nearrow)$ causes the pattern $(\searrow,\nearrow)$).

After identifying the type of causality, we must quantify the causal strength. The causal strength is measured by the accuracy of cross estimation between state spaces. In particular, we estimate the average pattern of $y(t)$ contemporaneously with $x(t)$ by a mapping that keeps the weights w_i^x on A_X, see Equations 6-9:

$$\widehat{y}(t_{x(i)}) = \left(Y(t_{x(i)}), Y(t_{x(i)}-\tau), \cdots, Y(t_{x(i)}-(E-1)\tau)\right), \widehat{y}(t_{x(i)}) \in \widehat{N}_{y(t)}, \quad (6)$$

$$h_i^{\widehat{y}} = \left(\frac{Y(t_{x(i)}-\tau) - Y(t_{x(i)})}{Y(t_{x(i)})}, \cdots, \frac{Y(t_{x(i)}-(E-1)\tau) - Y(t_{x(i)}-(E-2)\tau)}{Y(t_{x(i)}-(E-2)\tau)} \right), \quad (7)$$

$$\widehat{H}_{y(t)} = \sum_{i=1}^{E+1} w_i^x h_i^{\widehat{y}}, \quad (8)$$

$$\widehat{P}_{y(t)} = Pattern\left(\widehat{H}_{y(t)}\right). \quad (9)$$

By examining all state points on A_X and A_Y, we obtain all the estimated average patterns of $y(t)$ based on the dynamics of $x(t)$ and the real average patterns of $y(t)$. Then, accuracy can be used as a metric of causal strength where the estimated

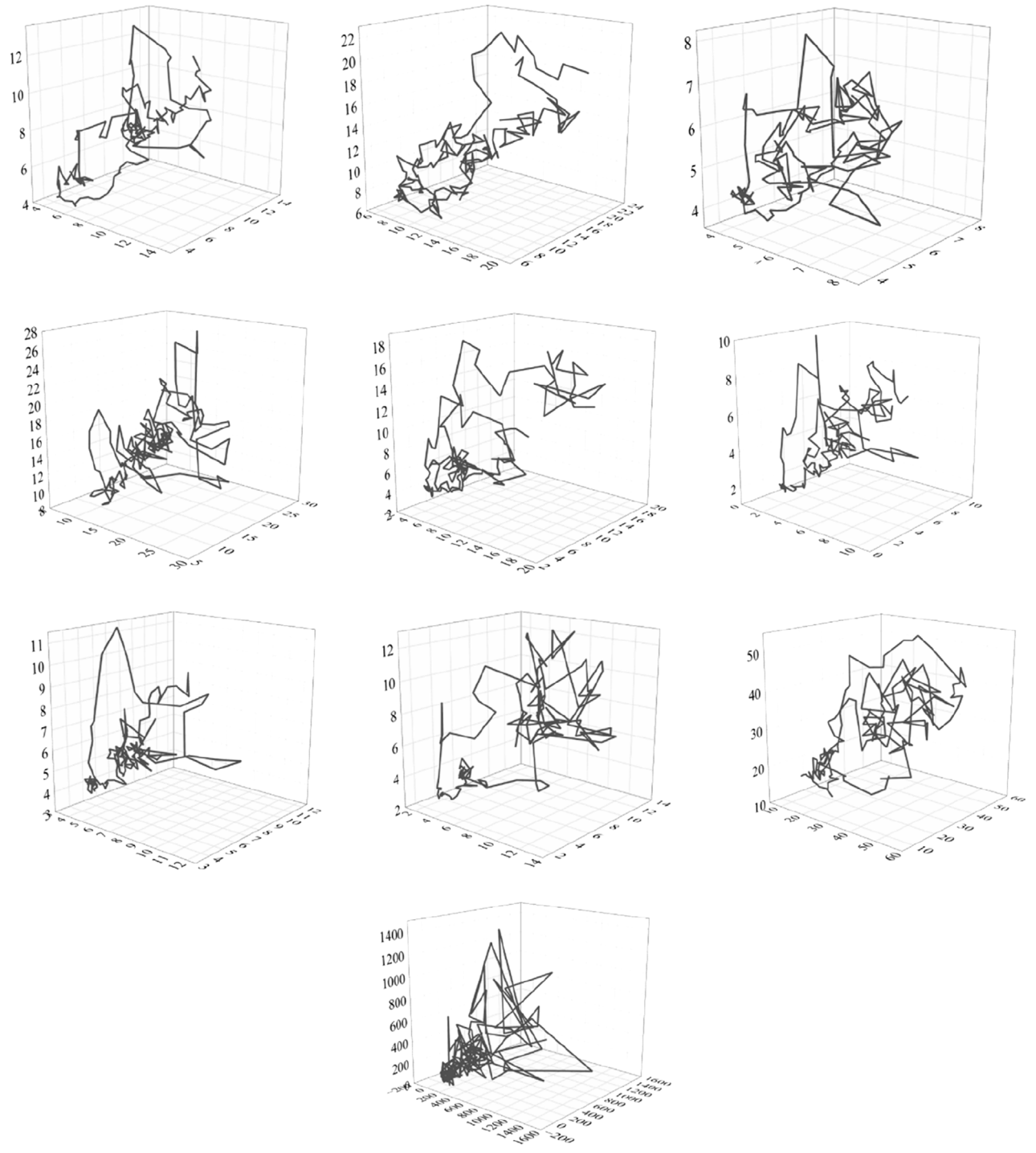


FIGURE 5
The reconstructed 3-dimensional state space based on the time series of sample variables. Note: the results are ordered based on Figure 2.

average pattern of $y(t)$ is equal to the real average pattern. Thus, for each possible pattern $P_{x(t)}$, we calculate the percentage of occasions for which the contemporaneously estimated average pattern $\widehat{P}_{y(t)}$ equals the true average pattern $P_{y(t)}$. Because there are three kinds of causalities, we can measure the causal strength for positive, negative, and dark causalities (more details can be seen in [40]).

3 Data

This work aims to detect the impact of economic policy on listed oil companies. Thus, we collect the sample data about economic policy and listed oil companies. For the economic policy, we use a widely used index, namely, the economic policy uncertainty index (EPU), which is a newspaper-based index for policy uncertainty

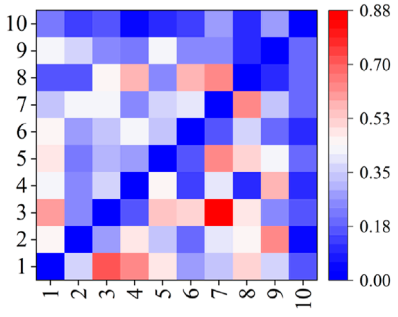


FIGURE 6
The positive causal impact among selected observations.

TABLE 2 The positive impacts among selected sample variables.

	1	2	3	4	5	6	7	8	9	10
1	0	0.47	0.59	0.42	0.47	0.46	0.34	0.15	0.44	0.21
2	0.36	0	0.25	0.26	0.22	0.28	0.43	0.15	0.38	0.14
3	0.72	0.29	0	0.38	0.3	0.33	0.41	0.46	0.25	0.18
4	0.63	0.5	0.16	0	0.29	0.44	0.26	0.57	0.21	0.04
5	0.48	0.34	0.53	0.44	0	0.33	0.36	0.25	0.42	0.11
6	0.26	0.18	0.5	0.15	0.15	0	0.39	0.58	0.24	0.13
7	0.35	0.4	0.88	0.41	0.63	0.17	0	0.63	0.25	0.28
8	0.5	0.46	0.48	0.11	0.5	0.35	0.63	0	0.11	0.11
9	0.37	0.63	0.25	0.58	0.42	0.18	0.33	0.11	0	0.28
10	0.17	0.06	0.15	0.12	0.2	0.1	0.19	0.18	0.19	0

Note the numbers from 1 to 10 represent the corresponding observations (see Table 1).

in China (More details are given at https://www.policyuncertainty.com/china_monthly.html). The monthly series from 2010/6/1 to 2025/2/1 are used. The length of the observed series is 177.

We consider the ten most popular listed oil companies: China National Petroleum Corporation, China Petrochemical Corporation, China National Offshore Oil Corporation, China Oilfield Services Limited, COSCO SHIPPING Energy Transportation, China Merchants Group, China Merchants Energy Transportation, Shenergy Company Limited, Guanghui Energy, and Jereh Group. Their monthly stock price series from 2010/6/1 to 2025/2/1 are used as sample. Their corresponding stock codes are given in Table 2. All the datasets are free at https://cn.investing.com/. The datasets are depicted in Figure 2.

From the statistical analysis, the stock price of China National Petroleum Corporation shows relatively higher fluctuations with a standard deviation of 20.3. Meanwhile, the stock price for China National Offshore Oil Corporation remains stable and has the lowest standard deviation. It is obvious the EPU has the highest volatility due to the highest standard deviation (see Table 1). Based on the correlation analysis, many stock prices show positive connections, while negative correlations are discovered between the EPU index and stock prices of other listed oil companies; see Figure 3.

4 Results

4.1 Parameter selections and dynamics reconstructions

According to our dynamical causal interaction methods, we need to detect dynamical impacts from the reconstructed state space. On this basis, two fundamental parameters must be identified, that is, embedding dimension E and time lag τ. Parameter E determines the goal dimension of reconstructed space, and time lag τ is the basis of the coordinates of reconstructed space. Several classic approaches are proposed to identify these parameters. In this work, we employ a widely used method to find the optimal embedding dimension, that is, the false nearest neighbor [49]. The main idea is to examine how the number of neighbors of a point along a trajectory changes with increasing embedding dimension. For a too-low embedding dimension, many of the neighbors will be false, but in an appropriate embedding dimension or higher, the neighbors are real. With increasing dimensions, the false neighbors will no longer be neighbors. Therefore, it is possible to find a faithful embedding dimension by examining how the number of neighbors changes as a function of dimension. Thus, when we reduce the dimension by one, some points are strongly affected and become false nearest neighbors (FNN). To identify these FNN points, we compared the distances between points in the E-dimensional space with those in the $E+1$-dimensional space and calculated their ratio. Economic policy uncertainty (EPU) is the critical factor in our work, and China National Petroleum Corporation is representative of the petroleum and petrochemical industry. Thus, the EPU and stock price series of China National Petroleum Corporation are selected as a sample. We determine $E = 3$ in this work; see Figure 4.

The time lag is another key parameter to guarantee a reliable reconstruction. If the delay time is too short, the two coordinate components of the phase space vector are so close numerically that they are indistinguishable from each other, thus failing to provide two independent coordinate components. If the delay time is too large, the two coordinate components appear to be completely independent, and there is no correlation between the projections of the trajectories of the chaotic attractor in the two directions. Therefore, a suitable method is needed to determine a suitable delay time to strike a balance between independence and correlation. This parameter is identified by mutual information that the first minimum in a plot between time delay and mutual information is a reliable selection [50]. We set $\tau = 5$ in this work.

Once we identify the embedding dimension and time delay, we can perform phase space reconstruction from each time series. The reconstructed structures are given in Figure 5. In this way, we map each one-dimensional time series into a 3-dimensional state space. The trajectory depicts the dynamical evolutions for each variable. On this basis, it is possible to identify the dynamic interactions between variables.

TABLE 3 The net positive causal impacts among samples.

	1	2	3	4	5	6	7	8	9	10
1	0	0.11	−0.13	−0.21	−0.01	0.2	−0.01	−0.35	0.07	0.04
2	−0.11	0	−0.04	−0.24	−0.12	0.1	0.03	−0.31	−0.25	0.08
3	0.13	0.04	0	0.22	−0.23	−0.17	−0.47	−0.02	0	0.03
4	0.21	0.24	−0.22	0	−0.15	0.29	−0.15	0.46	−0.37	−0.08
5	0.01	0.12	0.23	0.15	0	0.18	−0.27	−0.25	0	−0.09
6	−0.2	−0.1	0.17	−0.29	−0.18	0	0.22	0.23	0.06	0.03
7	0.01	−0.03	0.47	0.15	0.27	−0.22	0	0	−0.08	0.09
8	0.35	0.31	0.02	−0.46	0.25	−0.23	0	0	0	−0.07
9	−0.07	0.25	0	0.37	0	−0.06	0.08	0	0	0.09
10	−0.04	−0.08	−0.03	0.08	0.09	−0.03	−0.09	0.07	−0.09	0

TABLE 4 The negative impacts among selected sample variables.

	1	2	3	4	5	6	7	8	9	10
1	0	0.21	0.1	0.33	0.21	0.1	0.33	0.31	0.29	0.34
2	0.21	0	0.17	0.5	0.29	0.15	0.45	0.03	0.3	0.5
3	0.25	0.17	0	0.33	0.44	0.54	0.11	0.19	0.14	0.18
4	0.33	0.25	0.21	0	0.25	0.3	0.15	0.12	0.71	0.13
5	0.15	0.15	0.06	0.06	0	0.11	0.17	0.04	0.19	0.3
6	0.14	0.41	0.17	0.1	0.1	0	0.16	0.07	0.11	0.13
7	0.33	0.22	0.15	0.11	0.17	0.14	0	0.13	0.14	0.21
8	0.11	0.03	0.03	0.1	0.44	0.33	0.08	0	0.08	0.31
9	0.54	0.11	0.11	0.14	0.09	0.44	0.1	0.46	0	0.12
10	0.55	0.38	0.13	0.33	0.4	0.68	0.78	0.23	0.63	0

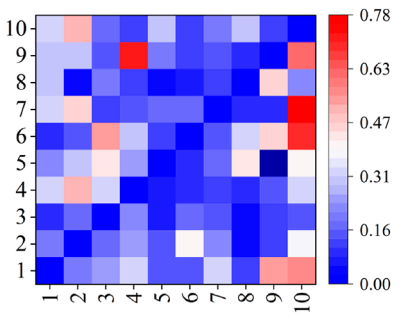


FIGURE 7
The negative causal impact among selected observations.

4.2 The positive causal impacts of economic policy on listed oil companies

In this section, we discuss the complex impacts of economic policy on the stock prices of oil companies. Causality analysis is a useful approach to reveal the causal impact between observations. We employ pattern causality to reveal their complex causal interactions. According to pattern causality, three kinds of causal interactions may exist: positive, negative, and dark causality. We first examine the positive causality between economic policy and the stock prices of listed oil companies. For each observation pair, the positive causal interactions are quantified by measuring the consistent patterns between state points on the reconstructed state space (see Figure 6). Table 2 depicts the pair-wise causal interactions for all sample variables. The direction of causality is from row variables to column variables. For example, the value 0.47 in the first row and second column (1,2) represents the causal interaction from variable 1 to variable 2, and the causal strength is 0.47. The results show positive causal interactions between sample variables. The average positive causal interaction equals 0.3. In particular, the average positive causal interaction between the listed oil companies is 0.34, while it is 0.16 between EPU and listed oil companies. These findings suggest that stock prices of listed oil companies are more likely to evolve in the same direction, and EPU also has relatively low positive causal impacts on the stock prices of listed oil companies. However, stock prices of oil companies also generate positive feedback to EPU. Moreover, the positive causal impacts are asymmetric. See Figure 6 and Table 3, which report the difference between received causal impacts and output impacts.

4.3 The negative causal impacts of economic policy on listed oil companies

In addition to positive causal impacts, we then examine whether there exist negative impacts between EPU and stock prices of

TABLE 5 The net negative causal impacts among samples.

	1	2	3	4	5	6	7	8	9	10
1	0	0	−0.15	0	0.06	−0.04	0	0.2	−0.25	−0.21
2	0	0	0	0.25	0.14	−0.26	0.23	0	0.19	0.12
3	0.15	0	0	0.12	0.38	0.37	−0.04	0.16	0.03	0.05
4	0	−0.25	−0.12	0	0.19	0.2	0.04	0.02	0.57	−0.2
5	−0.06	−0.14	−0.38	−0.19	0	0.01	0	−0.4	0.1	−0.1
6	0.04	0.26	−0.37	−0.2	−0.01	0	0.02	−0.26	−0.33	−0.55
7	0	−0.23	0.04	−0.04	0	−0.02	0	0.05	0.04	−0.57
8	−0.2	0	−0.16	−0.02	0.4	0.26	−0.05	0	−0.38	0.08
9	0.25	−0.19	−0.03	−0.57	−0.1	0.33	−0.04	0.38	0	−0.51
10	0.21	−0.12	−0.05	0.2	0.1	0.55	0.57	−0.08	0.51	0

TABLE 6 The dark impacts among selected sample variables.

	1	2	3	4	5	6	7	8	9	10
1	0	0.4	0.24	0.5	0.3	0.15	0.53	0.58	0.25	0.07
2	0.38	0	0.28	0.18	0.31	0.31	0.23	0.24	0.5	0.38
3	0.58	0.21	0	0.61	0.17	0.28	0.48	0.28	0.13	0.25
4	0.31	0.19	0.53	0	0.41	0.34	0.6	0.34	0.23	0.45
5	0.55	0.19	0.22	0.11	0	0.31	0.27	0.08	0.08	0.29
6	0.35	0.23	0.24	0.23	0.39	0	0.34	0.24	0.29	0.13
7	0.21	0.29	0.77	0.24	0.25	0.21	0	0.65	0.15	0.14
8	0.54	0.33	0.53	0.2	0.76	0.39	0.23	0	0.29	0.2
9	0.28	0.69	0.25	0.19	0.52	0.35	0.15	0.34	0	0.33
10	0.42	0.33	0.38	0.39	0.77	0.46	0.43	0.59	0.68	0

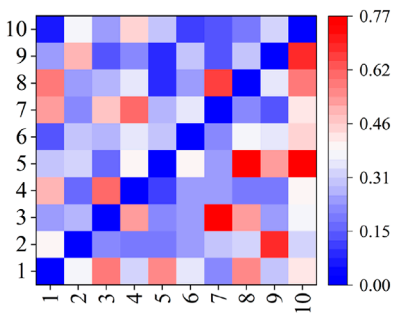


FIGURE 8
The dark causal impact among selected observations.

listed oil companies. According to pattern causality, the negative causal interactions can be quantified by measuring the opposite pattern of state points on the reconstructed state space. The results show negative causality between selected samples (Table 4). In total, the average negative causal impacts among all observations and only listed oil companies are 0.21 and 0.18. The average negative causal impact between EPU and listed oil companies is 0.33. These findings suggest that the negative causal impacts between the stock prices of listed oil companies are relatively lower than the impact between economic policy and listed oil companies.

To further analyze the causal impacts, it is obvious that the received impacts and output impacts are not equal. This indicates asymmetric negative causal interactions between samples (see Figure 7). In addition, we examine the net causal impact from the difference between received and output impacts (see Table 5). The results show that the EPU is more likely to bring net negative impacts on the stock prices of listed oil companies. In general, EPU has higher negative impacts on the stock prices of listed oil companies than negative feedback from oil companies.

4.4 The dark causal impacts of economic policy on listed oil companies

At this point, we have revealed positive and negative causal interactions between economic policy and stock prices of listed oil companies. Both positive and negative causality are widely reported by many previous works. However, we further detect whether a more complex causal interaction exists whose causal impact natures are quite different from negative and positive causality, namely, dark causality. According to pattern causality methods, the dark causal interaction can be measured by quantifying the complex patterns (neither consistent nor opposite) of state points in the reconstructed state space. The results report the dark causal

TABLE 7 The net dark causal impacts among samples.

	1	2	3	4	5	6	7	8	9	10
1	0	0.02	−0.34	0.19	−0.25	−0.2	0.32	0.04	−0.03	−0.35
2	−0.02	0	0.07	−0.01	0.12	0.08	−0.06	−0.09	−0.19	0.05
3	0.34	−0.07	0	0.08	−0.05	0.04	−0.29	−0.25	−0.12	−0.13
4	−0.19	0.01	−0.08	0	0.3	0.11	0.36	0.14	0.04	0.06
5	0.25	−0.12	0.05	−0.3	0	−0.08	0.02	−0.68	−0.44	−0.48
6	0.2	−0.08	−0.04	−0.11	0.08	0	0.13	−0.15	−0.06	−0.33
7	−0.32	0.06	0.29	−0.36	−0.02	−0.13	0	0.42	0	−0.29
8	−0.04	0.09	0.25	−0.14	0.68	0.15	−0.42	0	−0.05	−0.39
9	0.03	0.19	0.12	−0.04	0.44	0.06	0	0.05	0	−0.35
10	0.35	−0.05	0.13	−0.06	0.48	0.33	0.29	0.39	0.35	0

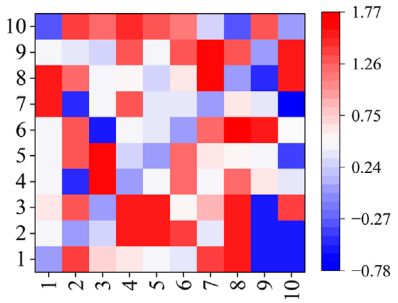


FIGURE 9
The dominant causal impacts in our samples. Note: negative values represent the negative causal impacts. Positive causal impacts are given by the values from [0,1]. The values larger than 1 represent the dark values. To clarify dark and positive causality, we add 1 to the original strength of dark causality.

impact between economic policy and the stock prices of listed oil companies (see Table 6). The average dark causal impact is 0.31. The average dark causal impact is 0.3 among the stock prices of listed oil companies, and the average dark causal impact approaches 0.45 between EPU and stock prices of listed oil companies. We can find relatively higher dark causal impacts between economic policy and listed oil companies than the impacts within listed oil companies.

Moreover, the dark causal interactions between economic policy and listed oil companies are also asymmetric (see Figure 8). We further calculate the difference between received and output impacts; see Table 7. It is clear that the output dark impacts from economic policy to listed oil companies are higher than the received impacts. This suggests that EPUs are more likely to affect the stock prices of listed oil companies in a complex way.

4.5 The dominant causal impacts of the economic policy on listed oil companies

We have revealed positive, negative, and dark causal impacts between economic policy and listed oil companies. The three different causal impacts show different magnitudes of impacts. Thus, it is interesting to compare their strength and find the dominant causal impacts. To this end, we compare three kinds of causal impacts between each variable pair and select the highest causal interaction. The results are given in Figure 9. Dark causality plays a dominant role in our samples. Within listed oil companies, positive causality occupies a relatively higher proportion, while dark causality occupies dominant roles between economic policy and listed oil companies. These findings suggest that the causal impacts between economic policy and listed oil companies are complex, neither purely positive nor purely negative, as reported by many previous methods.

The simultaneous existence of three different causal types between EPU and listed crude oil companies suggests a complex nonlinear impact relationship between them. Second, the strength of the three causal relationships is asymmetric, with dark causality dominating, indicating that the association between EPU and listed crude oil companies is volatile and complex. These findings help us understand the relationship between them. In addition, this complex influence relationship may make the economic system more risky, which provides the relevant market investors and market regulators with corresponding decision support.

5 Conclusion and discussion

This work employs the pattern causality method to reveal complex causality between economic policy and listed oil companies. Unlike previous findings that only one kind of causal impact between them, our study suggests that three different kinds of causal impacts simultaneously exist: positive, negative, and dark

causality. Our work may have several implications for practical application and theoretical analysis.

i) Three different kinds of causal impacts between economic policy and listed oil companies suggest more complex impacts between them, which gives new inspiration to understanding complex associations between them.
ii) The causal impacts of economic policy uncertainty on stock prices of listed oil companies are dominated by dark causality, which suggests neither the same nor opposite fluctuation between them. This may give a new perspective for management. In particular, dark causality denotes more complex connections. The underlying risk may be underestimated, and market participants should reconsider risk.
iii) Market players should upgrade risk assessment methods to capture "dark causality" risks. Oil firms and investors need more advanced models, as traditional ones overlook hidden risks. Firms should use real-time policy analysis tools and conduct stress tests. Regulators should enforce clear risk disclosures about such uncertainties. To manage these underestimated risks, oil companies should diversify operations geographically and across products. Investors should diversify portfolios, too. Both should create contingency plans for rapid response to policy changes, accounting for the unpredictable nature of dark causality.

As discussed in this work, dark causality may drive two variables to fluctuate in a nonlinear manner, moving in neither perfect unison nor in complete opposition, indicating a more intricate relationship between them. The underlying economic implications of dark causality can be understood from the following aspects:

First, dark causality offers a more realistic explanation for phenomena observed in real-world economic systems. In these systems, the fluctuations between variables rarely exhibit simple, unidirectional patterns. Instead, they often display complex, intertwined trends that deviate from straightforward positive or negative correlations.

Second, it provides novel evidence for the inherent complexity of economic systems. By uncovering these hidden causal relationships, we gain deeper insights into the intricate mechanisms that govern economic behavior, enhancing our understanding of how economic systems operate and evolve.

Third, dark causality reveals an even more convoluted nature of the relationships among economic variables. This complexity gives rise to increased uncertainty, and when dark causality dominates, it may signal elevated risks. As such, dark causality has the potential to be linked to systemic risks within economic systems.

Despite several interesting findings being discovered, there are also some limitations that can be solved in the future. In particular, dark causality is a newly defined interaction, and its underlying explanation needs to be discussed further. In addition, it is also difficult to distinguish impacts from other external factors, such as other related markets.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding authors.

Author contributions

YD: Funding acquisition, Investigation, Resources, Writing – review and editing, Methodology, Formal Analysis, Project administration, Conceptualization. XJ: Validation, Writing – review and editing, Supervision, Conceptualization. HX: Funding acquisition, Conceptualization, Formal Analysis, Writing – review and editing, Validation. ZW: Writing – review and editing, Formal Analysis, Supervision, Investigation, Project administration. XZ: Methodology, Data curation, Writing – review and editing, Conceptualization, Investigation. YW: Validation, Writing – review and editing, Writing – original draft, Data curation, Software, Visualization.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

Authors YD, XZ, XJ, and HX were employed by PetroChina Southwest Oil & Gasfield Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Dafermos Y, Nikolaidi M, Galanis G. Climate change, financial stability and monetary policy. *Ecol Econ* (2018) 152:219–34. doi:10.1016/j.ecolecon.2018.05.011

2. Danisman G, Bilyay-Erdogan S, Demir E. Economic uncertainty and climate change exposure. *J Environ Manage* (2025) 373:123760. doi:10.1016/j.jenvman.2024.123760

3. Adra S, Gao Y, Huang J, Yuan J. Geopolitical risk and corporate payout policy. *Int Rev Financial Anal* (2023) 87:102613. doi:10.1016/j.irfa.2023.102613

4. Apergis N, Fahmy H. Geopolitical risk and energy price crash risk. *Energy Econ* (2024) 140:107975. doi:10.1016/j.eneco.2024.107975

5. Liu Y, Suo X, Du X, Wu H, Lin H. Corporate digital innovation and stock price crash risk. *Finance Res Lett* (2024) 66:105690. doi:10.1016/j.frl.2024.105690

6. Liu C, Xu J. Risk spillover effects of new global energy listed companies from the time-frequency perspective. *Energy* (2024) 292:130502. doi:10.1016/j.energy.2024.130502

7. Su F, Song N, Shang H, Fahad S. The impact of economic policy uncertainty on corporate social responsibility: a new evidence from food industry in China. *Plos One* (2022) 17:e0269165. doi:10.1371/journal.pone.0269165

8. Mitchell JV, Mitchell B. Structural crisis in the oil and gas industry. *Energy Policy* (2014) 64:36–42. doi:10.1016/j.enpol.2013.07.094

9. Silvestre H, Gomes RC, Lamba JR, Correia AM. Implementation of Brazil's energy policy through the national oil company: from institutional chaos to strategic order. *Energy Policy* (2018) 119:87–96. doi:10.1016/j.enpol.2018.04.028

10. Diaz-Rainey I, Gehricke SA, Roberts H, Zhang R. Trump vs. Paris: the impact of climate policy on US listed oil and gas firm returns and volatility. *Int Rev Financial Anal* (2021) 76:101746. doi:10.1016/j.irfa.2021.101746

11. Al-Thaqeb SA, Algharabali BG, Alabdulghafour KT. The pandemic and economic policy uncertainty. *Int J Finance Econ* (2022) 27:2784–94.

12. Lin BQ, Bai R. Oil prices and economic policy uncertainty: evidence from global, oil importers, and exporters' perspective. *Res In Int Business Finance* (2021) 56:101357. doi:10.1016/j.ribaf.2020.101357

13. Shafiullah M, Miah MD, Alam MS, Atif M. Does economic policy uncertainty affect renewable energy consumption? *Renew Energy* (2021) 179:1500–21. doi:10.1016/j.renene.2021.07.092

14. Apostolakis GN, Floros C, Gkillas K, Wohar M. Financial stress, economic policy uncertainty, and oil price uncertainty. *Energy Econ* (2021) 104:105686. doi:10.1016/j.eneco.2021.105686

15. Xiao XY, Tian Q, Hou S. Economic policy uncertainty and grain futures price volatility: evidence from China. *China Agric Econ Rev* (2019) 11:642–54. doi:10.1108/caer-11-2018-0224

16. Jiang W, Dong L, Liu X, Zou L. Volatility spillovers among economic policy uncertainty, energy and carbon markets-The quantile time-frequency perspective. *Energy* (2024) 307:132683. doi:10.1016/j.energy.2024.132683

17. Feng XG, Luo WJ, Wang Y. Economic policy uncertainty and firm performance: evidence from China. *J Asia Pac Economy* (2023) 28:1476–93. doi:10.1080/13547860.2021.1962643

18. Xu ZX. Economic policy uncertainty, cost of capital, and corporate innovation. *J Banking Finance* (2020) 111:105698. doi:10.1016/j.jbankfin.2019.105698

19. Liu GC, Zhang CS. Economic policy uncertainty and firms' investment and financing decisions in China. *China Econ Rev* (2020) 63:101279. doi:10.1016/j.chieco.2019.02.007

20. Che M, Wang L, &Li YJ. Global economic policy uncertainty and oil price uncertainty: which is more important for global economic activity? *Energy* (2024) 310:133305. doi:10.1016/j.energy.2024.133305

21. Liu XJ, Wang Y, Du W, Ma Y. Economic policy uncertainty, oil price volatility and stock market returns: evidence from a nonlinear model. *North Am J Econ Finance* (2022) 62:101777. doi:10.1016/j.najef.2022.101777

22. Ilyas M, Khan A, Nadeem M, Suleman MT. Economic policy uncertainty, oil price shocks and corporate investment: evidence from the oil industry. *Energy Econ* (2021) 97:105193. doi:10.1016/j.eneco.2021.105193

23. Bloom N. The impact of uncertainty shocks. *Econometrica* (2009) 77:623–85.

24. Belderbos R, Tong TW, Wu SB. Multinational investment and the value of growth options: alignment of incremental strategy to environmental uncertainty. *Strateg Management J* (2019) 40:127–52. doi:10.1002/smj.2969

25. Chen Y, Dong S, Qian S, Chung K. Impact of oil price volatility and economic policy uncertainty on business investment - insights from the energy sector. *Heliyon* (2024) 10:e26533. doi:10.1016/j.heliyon.2024.e26533

26. Lee C, Lee C, Li Y. Assessing oil price volatility co-movement with stock market volatility through quantile regression approach. *Resour Policy* (2021) 55:101309.

27. Song Y, Chen B, Wang XY, Wang PP. Defending global oil price security: based on the perspective of uncertainty risk. *Energy Strategy Rev* (2022) 41:100858. doi:10.1016/j.esr.2022.100858

28. Adelaja AO, Akaeze HO. Supply response, economic diversification and recovery strategy in the oil sector. *Energy Strategy Rev* (2018) 21:111–20. doi:10.1016/j.esr.2018.05.002

29. Liu LP, Lü Z, Yoon SM. Impact of policy uncertainty on stock market volatility in the China's low-carbon economy. *Energy Econ* (2025) 141:108056. doi:10.1016/j.eneco.2024.108056

30. Zheng TC, Wang B, Rajaeifar MA, Heidrich O, Zheng J, Liang Y, et al. How government policies can make waste cooking oil-to-biodiesel supply chains more efficient and sustainable. *J Clean Prod* (2020) 263:121494. doi:10.1016/j.jclepro.2020.121494

31. Rong G, Qamruzzaman M. Symmetric and asymmetric nexus between economic policy uncertainty, oil price, and renewable energy consumption in the United States, China, India, Japan, and South Korea: does technological innovation influence? *Frontier Energy Res* (2022) 10:973557. doi:10.3389/fenrg.2022.973557

32. Fu LL, Yuan DY, Teng JM. Asymmetric dynamic linkage between consumer sentiment, inflation expectations, and international energy prices: evidence from time-frequency wavelet and nonlinear analysis. *Plos One* (2024) 19:e0308097. doi:10.1371/journal.pone.0308097

33. Liu ZH, Zhang H, Ding Z, Lv T, Wang X, Wang D. When are the effects of economic policy uncertainty on oil-stock correlations larger? Evidence from a regime-switching analysis. *Econ Model* (2022) 114:105941. doi:10.1016/j.econmod.2022.105941

34. Song DM, Tumminello M, Zhou WX, Mantegna RN. Evolution of worldwide stock markets, correlation structure, and correlation-based graphs. *Physic Rev E* (2011) 84:026108. doi:10.1103/physreve.84.026108

35. Bai XW. Tanker freight rates and economic policy uncertainty: a wavelet-based copula approach. *Energy* (2021) 235:121383. doi:10.1016/j.energy.2021.121383

36. Zhao L. Global economic policy uncertainty and oil futures volatility prediction. *Finance Res Lett* (2023) 54:103693. doi:10.1016/j.frl.2023.103693

37. Rehman MU. Do oil shocks predict economic policy uncertainty? *Physica A* (2018) 498:123–36. doi:10.1016/j.physa.2017.12.133

38. Cao GX, Zhang Q, Li QC. Causal relationship between the global foreign exchange market based on complex networks and entropy theory. *Chaos Solitons and Fractals* (2017) 99:36–44. doi:10.1016/j.chaos.2017.03.039

39. Yang L, Hamori S. Systemic risk and economic policy uncertainty: international evidence from the crude oil market. *Econ Anal Policy* (2021) 69:142–58. doi:10.1016/j.eap.2020.12.001

40. Wu T, Gao X, An S, Liu S. Diverse causality inference in foreign exchange markets. *Int J Bifurcation Chaos* (2021) 31:2150070. doi:10.1142/s021812742150070x

41. Vyrost T, Lyocsa S, Baumohl E. Granger causality stock market networks: temporal proximity and preferential attachment. *Physica A* (2015) 427:262–76. doi:10.1016/j.physa.2015.02.017

42. Dash SR, Maitra D. Do oil and gas prices influence economic policy uncertainty differently: multi-country evidence using time-frequency approach. *Q Rev Econ Finance* (2021) 81:397–420. doi:10.1016/j.qref.2021.06.012

43. Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys Rev Lett* (2009) 103:238701. doi:10.1103/physrevlett.103.238701

44. Sugihara G, May R, Ye H, Hsieh C, Deyle E, Fogarty M, et al. Detecting causality in complex ecosystems. *Science* (2012) 338:496–500. doi:10.1126/science.1227079

45. Stavroglou SK, Pantelous AA, Stanley HE, Zuev KM. Hidden interactions in financial markets. *PNAS* (2019) 22:10646–51. doi:10.1073/pnas.1819449116

46. Wang WX, Lai YC, Grebogi C. Data based identification and prediction of nonlinear and complex dynamical systems. *Phys Rep* (2016) 644:1–76. doi:10.1016/j.physrep.2016.06.004

47. Wu T, An F, Gao X, Zhong W, Kurths J. A novel framework for direct multistep prediction in complex systems. *Nonlinear Dyn* (2023) 111(10):9289–304. doi:10.1007/s11071-023-08360-7

48. Takens F. Detecting strange attractors in turbulence. *Mathematics* (1981) 898:366–81.

49. Krakovska A, Mezeiova K, Budacova H. Use of false nearest neighbours for selecting variables and embedding parameters for state space reconstruction. *J Complex Syst* (2015) 2015:1–12. doi:10.1155/2015/932750

50. Wu T, Gao X, An F, Sun X, An H, Su Z, et al. Predicting multiple observations in complex systems through low-dimensional embeddings. *Nat Commun* (2024) 15:2242. doi:10.1038/s41467-024-46598-w

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 11 July 2025
DOI 10.3389/fphy.2025.1642788

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Lu Zhang,
Northern Illinois University, United States
Adrian Zoicas-Ienciu,
Universitatea Babes Bolyai, Romania

*CORRESPONDENCE
Linsen Yin,
yls@lixin.edu.cn

RECEIVED 07 June 2025
ACCEPTED 24 June 2025
PUBLISHED 11 July 2025

CITATION
Lv D, Zhou Y, Zhang J and Yin L (2025) Unintended consequences of quote limits policy on stock market dynamics: evidence from the MF-X-DMA analysis.
Front. Phys. 13:1642788.
doi: 10.3389/fphy.2025.1642788

COPYRIGHT
© 2025 Lv, Zhou, Zhang and Yin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Unintended consequences of quote limits policy on stock market dynamics: evidence from the MF-X-DMA analysis

Dayong Lv[1], Yaping Zhou[2], Jiarui Zhang[3] and Linsen Yin[1]*

[1]School of Financial Technology, Shanghai Lixin University of Accounting and Finance, Shanghai, China, [2]Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China, [3]School of Economics and Management, Tongji University, Shanghai, China

Introduction: This study investigates the effect of recently implemented quote limits policy in China's A-share Main Board market, a mechanism designed to enhance market stability. We examine whether this policy achieves its goals uniformly across distinct large-cap (HS300) and mid-to-small cap (CSI500) market segments.

Methods: Utilizing the Multifractal Detrending Moving Average Cross-Correlation analysis (MF-X-DMA) and the nonlinear Granger causality test, we assess the policy's potential heterogeneous impact on these two indices.

Results: Our findings reveal that the policy decreases market efficiency and increases cross-market co-movement complexity for HS300, while conversely improving market efficiency and information flow for CSI500.

Discussion: We offer novel empirical evidence on this policy using nonlinear methods. Our findings highlight the potentially divergent and unintended consequences of "one-size-fits-all" financial regulations across varied market segments, providing valuable insights for policymakers.

KEYWORDS

quote limits, MF-X-DMA, A-share, nonlinear Granger causality, cross-correlation

1 Introduction

Global financial markets, while important for capital allocation and economic growth, are often prone to extreme volatility and market anomalies. To mitigate systemic risks and safeguard investor interests, regulators worldwide have actively implemented various market stability mechanisms. For instance, in the United States, key mechanisms include comprehensive circuit breakers, order protection rule, and market-wide trading halts during periods of intense volatility. Similarly, Japan's equity markets utilize daily price range limits (often called "daily price limits") for individual stocks and futures contracts, alongside circuit breakers applied to major indices like the Nikkei 225, to ensure orderly price formation. These measures generally aim to curb excessive speculation, prevent "flash crashes," and ensure orderly trading. As the world's second-largest economy, China's financial market regulators have similarly intensified efforts to refine trading mechanisms, particularly in response to its burgeoning stock market and increasing participation.

A significant development in this regard is the recent introduction of the quote limits system (also called "price cage") in China's A-share Main Board market. This policy,

detailed in revised trading rules issued by the Shanghai Stock Exchange and Shenzhen Stock Exchange on 17 February 2023, became effective on 10 April 2023, coinciding with the listing of the first batch of registration-based main board stocks. Specifically, during the continuous auction phase (9:30-11:30 and 13:00-14:57), buy orders are restricted to a maximum of 102% of the benchmark selling price, while sell orders cannot be below 98% of the benchmark buying price. This "refusal mechanism" for invalid orders is designed to prevent large abnormal quotes, counter manipulative behaviors, and ultimately enhance market stability.

Despite the recent implementation and significant regulatory implications of this policy, empirical research on the A-share Main Board's quote limits remains limited. While existing literature [1–5] extensively analyzes the impact of broader price limit policies (e.g., daily price limits or circuit breakers) on market quality, few studies have investigated the effects of the more continuously applied quote limits, particularly within the unique context of China's market and on distinct market segments. This research gap presents an important opportunity to contribute to a deeper understanding of real-world policy outcomes.

To understand whether this policy achieves its intended goals uniformly across diverse market segments, this study specifically focuses on two pivotal and distinct benchmarks of the Chinese equity market: the large-cap HS 300 (HS300) and the mid-to-small cap CSI 500 (CSI500) indices. Given that component stocks of these two indices differ in market capitalization, investor composition, and liquidity profiles, the quote limits policy may exert differential impacts [6–10]. For instance, CSI500 stocks are often more susceptible to manipulative practices, as their relatively smaller market size allows for significant price fluctuations with comparatively less capital. Conversely, extreme price movements in large-cap stocks (HS300) are less likely to be driven purely by manipulation and may, instead, more frequently stem from informed trading activities or the rapid incorporation of significant news. Therefore, we hypothesize that the quote limits policy might differentially impact these two market segments: for the HS300, it could inadvertently restrict the execution of legitimate informed trades, thereby potentially distorting information processing and leading to decreased market efficiency. Conversely, for the mid-to-small cap CSI500, the policy might effectively curb speculative and manipulative behaviors, thus fostering greater market orderliness and enhancing overall efficiency.

Understanding financial market dynamics requires moving beyond traditional linear frameworks, as market returns often exhibit complex, non-Gaussian, and multi-scale characteristics [11–19]. For example, Gu and Huang [20] and Zhou et al. [21] demonstrate that daily stock returns frequently display high kurtosis (indicating "fat tails") and non-zero skewness, highlighting their significant deviations from normality. Such widely observed features, including leptokurtosis and long-range correlations, are very common in financial time series [17, 22]. Traditional linear models are often insufficient to capture these intricate aspects, such as varying degrees of market efficiency across different scales and complex nonlinear interdependencies [23]. Therefore, advanced non-linear and multifractal analysis tools are essential for comprehensively uncovering the underlying mechanisms and policy impacts [12, 21, 24].

A substantial body of literature, grounded in fractal theory and nonlinear dynamics, provids a suite of powerful tools for such analyses. Based on the fractal theory, some studies propose various multifractal analysis techniques, such as Detrended Fluctuation Analysis (DFA) [25], Multifractal Detrended Fluctuation Analysis (MF-DFA) [26], Detrended Cross-Correlation Analysis (DCCA) [27], Multifractal Detrended Cross-Correlation Analysis (MF-DCCA) [28], and Multifractal Detrending Moving Average Cross-Correlation Analysis (MF-X-DMA) [29]. These methods are also widely used to analyze the structure of financial markets and the cross-correlations among them. For example, Zhou et al. [21] and Acikgoz [16] utilize MF-DCCA to examine multifractal characteristics and long-range cross correlations among individual stocks or financial markets. Among these methods, the MF-X-DMA is known for its enhanced robustness in removing local trends, making it particularly well-suited for analyzing non-stationary financial time series [30]. Following the work of Jiang and Zhou [29], Li et al. [30] apply the MF-X-DMA method to test exchange rate regime reform. Mikhaylov et al. [31] use the MF-X-DMA to examine the factors affecting open innovation. Utilizing MF-X-DMA, Ruan et al. [1] show that relaxing daily price limits increases market multifractality. To complement the analysis of complex market structures, the nonlinear Granger causality test proposed by Baek and Brock [32] has become a widely-used tool for identifying directional influence in nonlinear financial systems [1, 15, 33]. For instance, Zhou et al. [34] use it to assess the effect of options trading. Zhang et al. [24] employ nonlinear Granger causality to map systemic risk networks. Vogl and Kojić [35] also use nonlinear Granger causality to test the interplay between cryptocurrencies and related indices.

Based on the implementation of the quote limits policy in the Chinese A-share Main Board market, we use MF-X-DMA and nonlinear Granger causality test to examine the impact of this policy on the HS300 and CSI500 indices. Our findings, consistent with the aforementioned hypothesized divergence, show that the policy reduces the market efficiency of the HS300 index, but improves the efficiency of the CSI500 index. Furthermore, in the cross-market context, the policy increases the complexity of co-movement and strengthens futures-to-spot causality for the HS300, which suggests potentially less efficient cross-market information transmission. In contrast, for the CSI500, it simplifies co-movement complexity and establishes a more integrated bidirectional nonlinear causality, indicating improved information flow.

This research contributes to the existing literature in several ways. First, this is among the first empirical studies to investigate the effects of the recently implemented and more granular quote limits policy on China's A-share Main Board. Unlike studies on broader price limits [1, 4, 5, 36], our focus on this precise mechanism and its effects on distinct market segments offers a more targeted understanding of contemporary Chinese market regulation.

Second, by extending the conventional event study framework, we employ a sophisticated combination of advanced methodologies, including MF-X-DMA and nonlinear Granger causality tests. These methods are uniquely suited to uncover the complex, non-linear, multifractal characteristics and causal relationships within financial time series [1, 15, 16, 20–22, 37, 38], which standard linear models might overlook. This rigorous approach provides a more

comprehensive assessment of market efficiency and information transmission dynamics under policy interventions.

Third, our findings offer empirical evidence of the heterogeneous impact of a single regulatory policy on different market capitalization segments (large-cap HS300 vs. mid-to-small cap CSI500). This highlights that a "one-size-fits-all" regulatory approach may lead to divergent, and possibly unintended, consequences across varied market ecosystems, offering valuable insights for policymakers designing tailored and effective regulations.

The remainder of this paper is organized as follows. Section 2 introduces the methodologies used in this paper. Section 3 details our data and provides descriptive statistics. Section 4 presents the analysis of the quote limits' effect on individual market efficiency, followed by Section 5, which examines its impact on spot-futures market cross-correlations and causality. Finally, Section 6 concludes with a summary of findings, policy implications, and future research directions.

2 Methodology

Given the multifractal and nonlinear nature of financial markets, this paper follows related literature [11, 13–15, 21, 38] and employs a combination of methodologies to investigate the multifaceted impact of China's Main Board quote limits policy on market efficiency and inter-market linkages. Specifically, we utilize the MF-X-DMA to uncover the multifractal characteristics of individual market returns and their cross-correlations, and the Nonlinear Granger Causality test to identify the directional causal relationships between spot and futures markets. This section provides a brief introduction to these two methods.

2.1 MF-X-DMA

Financial time series often exhibit complex, multi-scale behaviors that linear models cannot fully capture. The MF-X-DMA, originally proposed by Jiang and Zhou [29], is a robust technique well-suited for analyzing such non-stationary and multifractal properties, particularly in cross-correlated series. Its primary advantage lies in effectively removing local trends, thereby providing a more accurate estimation of scaling exponents compared to other methods like MF-DFA. Therefore, we also apply this method to both individual index returns (HS300 and CSI500) and their respective spot-futures cross-correlations. The core steps of the MF-X-DMA procedure are as follows:

Step 1: Constructing integrated profiles

For two time series, $X = \{x_t\}_{t=1}^{N}$ and $Y = \{y_t\}_{t=1}^{N}$, with length N, we first construct their integrated profiles, which represent the cumulative deviations from their respective means:

$$X(t) = \sum_{i=1}^{t} (x_i - \bar{x}), t = 1, 2, \ldots, N \tag{1}$$

$$Y(t) = \sum_{i=1}^{t} (y_i - \bar{y}), t = 1, 2, \ldots, N \tag{2}$$

where $\bar{x}$ and $\bar{y}$ denote the average values of the series x_t and y_t, respectively.

Step 2: Segmentation of profiles

The integrated profiles $X(t)$ and $Y(t)$ are divided into $N_s = int(N/s)$ non-overlapping segments, each of length s. To ensure complete coverage and robust analysis, this segmentation process is performed from both ends of the series, resulting in a total of $2N_s$ segments. The segment size s influences the scale of correlations detected, ranging from short-term (smaller s) to long-term (larger s).

Step 3: Detrending via moving average filtering

For each segment v (where v = 1,2, …,$2N_s$), and for both integrated profiles $X(t)$ and $Y(t)$, local trends are removed using a moving average filter. The moving average $Z(j)$ (where $Z \in \{X, Y\}$) within a window of size n is calculated based on a position parameter $\theta \in [0, 1]$ as follows:

$$Z_n(j) = \frac{1}{n} \sum_{k=-\lfloor (n-1)\theta \rfloor}^{\lfloor (n-1)(1-\theta) \rfloor} Z(j-k) \tag{3}$$

where $\lfloor g \rfloor$ denotes the largest integer not greater than g. The parameter θ determines the nature of the moving average: $\theta = 0$ corresponds to a backward moving average, $\theta = 0.5$ to a centered moving average, and $\theta = 1$ to a forward moving average. This step effectively detrends the series by subtracting the estimated local trend.

Step 4: Computing detrended covariance

The detrended covariance for each segment v is then calculated from the residuals after local trend removal as follows:

$$F^2(s, v) = \frac{1}{s} \sum_{j=1}^{s} \left[X_{(v-1)s+j} - \tilde{X}_v(j) \right] \left[Y_{(v-1)s+j} - \tilde{Y}_v(j) \right] \tag{4}$$

where, $\tilde{X}_v(j)$ and $\tilde{Y}_v(j)$ represent the fitting polynomials (of order m) for the integrated profiles $X(t)$ and $Y(t)$ within segment v. For the second set of segments (from the opposite end), a similar formula applies.

Step 5: Deriving the qth order fluctuation function

The overall qth order fluctuation function $F_q(s)$ is computed by averaging the detrended covariances across all segments as follows:

For $q \neq 0$:

$$F_q(s) = \left[\frac{1}{2N_s} \sum_{v=1}^{2N_s} (F^2(s, v))^{q/2} \right]^{1/q} \tag{5}$$

For $q = 0$:

$$F_0(s) = \exp \left[\frac{1}{2N_s} \sum_{v=1}^{2N_s} ln(F^2(s, v)) \right] \tag{6}$$

The parameter q allows for differentiation between small and large fluctuations: negative q values emphasize small fluctuations, while positive q values highlight large fluctuations.

Step 6: Determining the scaling exponent

The relationship between $F_q(s)$ and s typically follows a power law, i.e.:

$$F_q(s) \sim s^{H_{xy}(q)} \tag{7}$$

where $H_{xy}(q)$ is the generalized cross-correlation exponent. If $H_{xy}(q)$ varies with q, it indicates multifractality in the cross-correlation between $X(t)$ and $Y(t)$. A constant $H_{xy}(q)$ implies mono-fractal behavior. The persistence of cross-correlation is indicated by $H_{xy}(q) > 0.5$, while anti-persistence is suggested by $H_{xy}(q) < 0.5$. The series are uncorrelated if $H_{xy}(q) = 0.5$.

Step 7: Characterizing multifractality

To further quantify the multifractality, we derive the multifractal Renyi exponent $\tau_{xy}(q)$, the singularity strength $\alpha(q)$, and the multifractal spectrum $f(\alpha)$ using the Legendre transform as follows:

$$\tau_{xy}(q) = H_{xy}(q) - 1 \tag{8}$$

$$\alpha_{xy}(q) = H_{xy}(q) + qH'_{xy}(q) \tag{9}$$

$$f(\alpha) = q\left[\alpha(q) - H_{xy}(q)\right] + 1 \tag{10}$$

where $H'_{xy}(q)$ is the derivative of $H_{xy}(q)$ with respect to q.

The degree of multifractality is typically measured by (Equation 1) the width of the multifractal spectrum (i.e., $\Delta\alpha = max\left(\alpha_{xy}\right) - \min\left(\alpha_{xy}\right)$), and (Equation 2) the degree of multifractality (i.e., $\Delta h = H_{max}(q) - H_{min}(q)$). For a single time series, a larger $\Delta\alpha$ or Δh signifies a stronger degree of multifractality, implying lower market efficiency. In the context of cross-correlations between two markets, a larger $\Delta\alpha$ or Δh suggests a more complex and potentially riskier linkage.

2.2 Nonlinear granger causality test

While multifractal analysis reveals the complexity of market dynamics, understanding the directional influence between markets also requires a causality framework. Given the acknowledged nonlinear characteristics of financial data, traditional linear Granger causality tests may be insufficient. Therefore, we employ the non-linear Granger causality test, a non-parametric statistical method proposed by Baek and Brock [32], to analyze causal relationships, particularly between spot and futures markets for the HS300 and CSI500 indices. This method is important for uncovering intricate, non-linear predictive power that might be obscured by linear approaches.

The fundamental idea of the non-linear Granger causality test is to assess whether past values of one time series (Y) improve the prediction of another time series (X), beyond what can be predicted by past values of X alone, while considering non-linear dependencies.

Specifically, the nonlinear Granger causality test is conducted as follows:

Let $\{X_t\}$ and $\{Y_t\}$ be two stationary and weakly dependent time series of length N. We define lagged and lead vectors for these series. For instance, an α-length lag vector is $X_{t-\alpha}^{\alpha} = (X_{t-\alpha}, X_{t-\alpha+1}, \ldots, X_{t-1})$, and an m-length lead vector for X at time t is $X_t^m = (X_t, X_{t+1}, \ldots, X_{t+m-1})$. Similar definitions apply to Y.

The null hypothesis H_0 that Y does not non-linearly Granger cause X is formally expressed by comparing the following conditional probabilities:

$$\begin{aligned} &Pr\left(\left\|X_t^m - X_s^m\right\| < \varepsilon \middle| \left\|X_{t-\alpha}^{\alpha} - X_{s-\alpha}^{\alpha}\right\| < \varepsilon, \left\|Y_{t-\beta}^{\beta} - Y_{s-\beta}^{\beta}\right\| < \varepsilon\right) \\ &= Pr\left(\left\|X_t^m - X_s^m\right\| < \varepsilon \middle| \left\|X_{t-\alpha}^{\alpha} - X_{s-\alpha}^{\alpha}\right\| < \varepsilon\right) \end{aligned} \tag{11}$$

where, Pr $(\cdot)$ denotes probability, $\|\cdot\|$ is the maximum norm, and $\varepsilon > 0$ is a given distance. This equation essentially states that the probability of two m-length lead vectors of X being within ε distance given their respective pasts (both X and Y lags) is the same as given only X lags.

To operationalize this comparison, the conditional probabilities are expressed as ratios of joint probabilities, computed using correlation integral estimators. These estimators count the number of pairs (t,s) for which the specified vectors are within the distance ε. For example, a joint probability $C(m+\alpha, \beta, \varepsilon)$ counts pairs where the $(m+\alpha)$-length lead vector of X and the β-length lag vector of Y are simultaneously within ε distance.

The test statistic is then constructed based on these ratios of correlation integrals, and under the null hypothesis, it follows an asymptotically normal distribution as follows:

$$\sqrt{N}\left[\frac{C(m+\alpha,\beta,\varepsilon)}{C(\alpha,\beta,\varepsilon)} - \frac{C(m+\alpha,\varepsilon)}{C(\alpha,\varepsilon)}\right] \sim N(0,\sigma^2) \tag{12}$$

where σ^2 is the asymptotic variance. A statistically significant value of the test statistic leads to the rejection of the null hypothesis, indicating that Y non-linearly Granger causes X. We perform this test symmetrically to also examine if X nonlinearly Granger causes Y. This allows us to comprehensively assess the directionality and strength of information flow between the spot and futures markets of the HS300 and CSI500 indices before and after the implementation of quote limits.

3 Data, variables, and descriptive statistics

3.1 Data and variables

To investigate whether the implementation of quote limits has diverse impacts across different stock market segments, we select two prominent and distinct benchmarks of the Chinese equity market: the large-cap HS 300 Index (HS300) and the small-cap CSI 500 Index (CSI500). Specifically, the HS300 comprises 300 A-share stocks with large market capitalization and good liquidity listed on the Shanghai and Shenzhen stock exchanges, serving as a key indicator for the overall performance of China's large-cap, blue-chip companies. In contrast, the CSI500 includes 500 A-share stocks with medium and small market capitalization, representing the performance of mid-to-small cap enterprises and offering insights into a broader range of market participants and growth-oriented sectors.

Daily closing prices for both the spot indices (HS300 and CSI500) are sourced from the China Stock Market and Accounting Research (CSMAR) database, a widely used financial database in related literature [39–41]. To ensure stationarity and facilitate subsequent econometric and multifractal analyses, these price series

TABLE 1 Descriptive Statistics of daily returns for HS300 and CSI500.

Period	Mean	Median	Std	Kurt	Skew
Panel A: HS300					
Pre	−0.0004	−0.0004	0.0114	4.7300	−0.1061
Post	−0.0001	−0.0008	0.0116	15.3243	0.5949
Panel B: CSI500					
Pre	0.0001	0.0009	0.0114	6.6390	−0.9014
post	−0.0001	−0.0007	0.0150	13.9614	0.7104

Note: This table presents the descriptive statistics for daily logarithmic returns of HS300 and CSI500. "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after. "Mean," "Median," "Std.," "Kurt," and "Skew" are the mean, median, standard deviation, Kurtosis, and Skewness of daily returns.

are transformed into daily logarithmic returns, consistent with related literature [42–44]. Specifically, the daily logarithmic return (R_t) for an asset at day t is calculated as the difference between the natural logarithm of the current day's closing price (P_t) and that of the previous day's closing price (P_{t-1}), expressed as follows:

$$R_t = \ln(P_t) - \ln(P_{t-1}) \tag{13}$$

For a robust comparative analysis of the policy's effect, we use a specific observation window encompassing 2 years prior to and 2 years following the implementation of the quote limits policy. Specifically, the pre-event period ("Pre") is defined from 10 April 2021, to 9 April 2023, while the post-event period ("Post") extends from 10 April 2023, to 9 April 2025.

3.2 Descriptive statistics

Table 1 presents the descriptive statistics for the daily logarithmic returns of both the spot indices across these two distinct periods. As shown, the daily logarithmic returns for both HS300 and CSI500 indices consistently show low means, medians, and standard deviations across both pre- and post-event periods, indicating stable average returns and relatively consistent overall volatility. However, pronounced deviations from normality are clear in the higher-order moments. All series display kurtosis values greater than 3, indicating strong leptokurtosis or "fat tails." Notably, during the post-event period ("Post"), the kurtosis for HS300 and CSI500 returns dramatically increases, suggesting a heightened probability of extreme price movements. In addition, the skewness values are consistently non-zero, indicating asymmetry.

These non-normal distributions, particularly their heightened leptokurtosis and shifts in asymmetry following the implementation of quote limits, highlight the inherent multifractal nature of these financial time series. Therefore, it is necessary to employ sophisticated tools such as the MF-X-DMA approach to capture complex, scale-dependent correlations.

4 Effect of quote limits on stock market efficiency

4.1 Analysis of the generalized hurst exponent

We first use the MF-X-DMA method to calculate the generalized Hurst exponent, $H(q)$, for the HS300 and CSI500 index returns across for q ranging from −10 to 10. The results are presented in Table 2. As shown, for the HS300 index, $H(q)$ shows a noticeable decreasing trend with increasing q before the implementation of quote limits. This pattern signals persistence in small fluctuations and a shift towards anti-persistence in large fluctuations. After the event, persistence in small fluctuations strengthens, while the tendency for reversals after large price swings also becomes more pronounced. For instance, $H(q)$ at $q = -10$ rises sharply from 0.7978 to 0.8664, indicating strengthened persistence of small fluctuations. Conversely, the $H(q)$ value at $q = 10$ drops dramatically from 0.4029 to 0.1965, suggesting a much stronger tendency for price reversals following large swings.

In contrast, for the CSI500 index, $H(q)$ also decreases with increasing q before the event, showing persistence across most fluctuation sizes. After the implementation of quote limits, the persistence in small price movements weakens, while the tendency for reversals after large price swings strengthens, similar to the behavior of the HS300. For example, $H(q)$ at $q = -10$ decreases from 0.9013 to 0.7967, suggesting a weakening of small fluctuation persistence. Meanwhile, $H(q)$ at $q = 10$ drops from 0.3122 to 0.2445, indicating stronger anti-persistence for large fluctuations.

In sum, while both indices are multifractal and experience a stronger reversal tendency for large price swings post-event, the behavior of small price movements diverges significantly. The HS300 shows increased persistence in small fluctuations, whereas the CSI500 exhibits decreased persistence.

4.2 Change in the degree of multifractality

This section investigates the effect of quote limits on market efficiency by analyzing the change in the degree of multifractality in the HS300 and CSI500 index returns around the quote limits event. Specifically, we first follow related literature [1, 13, 14, 45, 46] and use two measures to proxy market efficiency: 1) the width of the multifractal spectrum ($\Delta\alpha$) and 2) the degree of multifractality (Δh). Accordingly, $\Delta\alpha$ measures the range of scaling exponents present in the data, indicating the heterogeneity of the fractal properties. Δh quantifies the extent to which a time series or dataset exhibits multifractal characteristics, implying that its scaling behavior requires a continuous spectrum of exponents for description, rather than a single fractal dimension. Consequently, larger $\Delta\alpha$ or Δh indicates stronger multifractality and thus lower market efficiency.

In addition, we follow related literature [47, 48] and use the Market Deficiency Measure (MDM) to inversely measure market efficiency. Specifically, MDM is computed as follows:

$$MDM = \frac{1}{2}(|H(-10) - 0.5| + |H(10) - 0.5|) \tag{14}$$

TABLE 2 Change in the Hurst exponent behavior of stock index return series.

q	HS300		CSI500	
	Pre	Post	Pre	Post
−10	0.7978	0.8664	0.9013	0.7967
−9	0.7866	0.8541	0.8893	0.7881
−8	0.7733	0.8393	0.8745	0.7779
−7	0.7570	0.8214	0.8561	0.7654
−6	0.7369	0.7994	0.8331	0.7499
−5	0.7114	0.7724	0.8041	0.7307
−4	0.6794	0.7403	0.7689	0.7072
−3	0.6422	0.7044	0.7287	0.6795
−2	0.6085	0.6687	0.6869	0.6494
−1	0.5866	0.6371	0.6464	0.6195
0	0.5719	0.6083	0.6075	0.5899
1	0.5567	0.5703	0.5685	0.5528
2	0.5376	0.5032	0.5273	0.4967
3	0.5151	0.4174	0.4848	0.4316
4	0.4920	0.3472	0.4447	0.3775
5	0.4704	0.2987	0.4099	0.3379
6	0.4517	0.2652	0.3813	0.3088
7	0.4359	0.2409	0.3583	0.2868
8	0.4228	0.2226	0.3397	0.2696
9	0.4120	0.2082	0.3246	0.2558
10	0.4029	0.1965	0.3122	0.2445

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after.

where H (-10) and H (10) are generalized Hurst exponents when $q = -10$ and $q = 10$, respectively; and $|\cdot|$ denotes the absolute value function. Similarly, higher *MDM* suggests stronger multifractality and potentially lower market efficiency due to increased predictability based on scale of fluctuations.

Table 3 reports results of changes in these measures for the HS300 and CSI500 indices. For the HS300 index, all three multifractal indicators ($\Delta\alpha$, Δh, and *MDM*) show a substantial increase after the implementation of the quote limits. This widening of the multifractal spectrum and the higher *MDM* value suggest a more pronounced multifractal structure and a greater disparity in the scaling behavior of different magnitude price fluctuations. Therefore, the increased multifractality implies that the implementation of quote limits likely led to a decrease in the market efficiency of the HS300 index.

TABLE 3 Change in multifractal spectra characteristics.

Period	α_{min}	α_{max}	$\triangle\alpha$	$\triangle h$	*MDM*
Panel A: HS300					
Pre	0.3119	0.9094	0.5975	0.3949	0.1975
Post	0.0801	0.9896	0.9095	0.6699	0.3349
Panel B: CSI500					
Pre	0.1880	1.0219	0.8339	0.5891	0.2946
Post	0.1312	0.8824	0.7512	0.5523	0.2761

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after.

In contrast, the CSI500 index exhibits a decrease in all three multifractal indicators ($\Delta\alpha$, Δh, and *MDM*) following the quote limits event. The narrower multifractal spectrum and the lower *MDM* value indicate a reduction in the strength of multifractality and a more homogeneous scaling behavior across different fluctuation magnitudes. This suggests that the quote limits event is associated with an increase in the market efficiency of the CSI500 index.

Figure 1 illustrates the changes in the multifractal spectrum, $f(\alpha)$, for both indices around the quote limits event. As depicted, the multifractal spectrum for the HS300 index becomes wider after the event, while that for the CSI500 index becomes narrower post-event. This also indicates that quote limits event strengthens multifractality in the HS300 index returns but attenuates multifractality in the CSI500 index returns.

In conclusion, the implementation of quote limits has a divergent impact on the market efficiency of the two indices. It decreases the market efficiency of the HS300 by increasing its multifractal characteristics, while conversely increasing the market efficiency of the CSI500 by reducing its multifractal characteristics.

4.3 Change in multifractal sources

This section investigates the potential impact of the quote limits event on the sources of multifractality in the HS300 and CSI500 index returns. According to Kantelhardt et al. [26], Zhou [12], and Kwapień et al. [49] multifractality in financial time series typically arises from two main sources: 1) a broad, non-Gaussian probability distribution of returns (e.g., fat tails) and 2) temporal correlations, which can be linear or nonlinear. Following established methodologies in related literature [1, 17, 22, 50, 51], we use surrogate analysis to disentangle these sources. Specifically, shuffling the original series destroys temporal correlations while preserving the amplitude distribution. In contrast, creating surrogate series preserves linear correlations but randomizes nonlinear structures. Therefore, significant multifractality remaining in the shuffled series suggests the importance of the amplitude distribution, while a reduction in multifractality in the surrogated series points to the role of nonlinear temporal correlations.

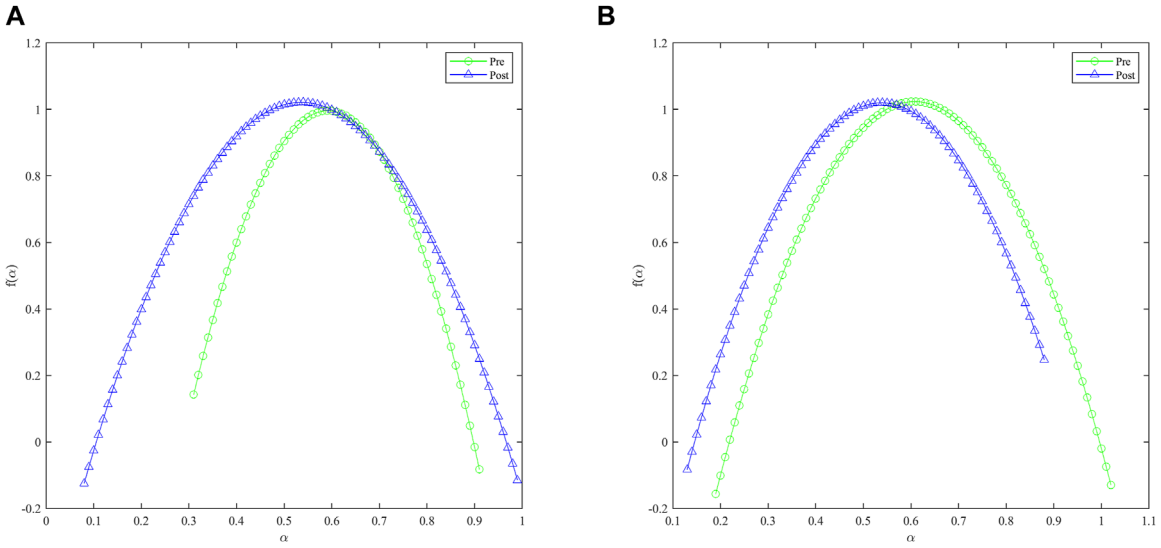


FIGURE 1
Change in Multifractal spectra. **(A)**: HS300. **(B)**: CSI500.

Table 4 reports the multifractal spectra characteristics of the shuffled and surrogated series for both the HS300 and CSI500 indices before and after the event. For both indices, there is a consistent reduction in multifractality of the shuffled and surrogated series both before and after the implementation of quote limits. Specifically, the shuffled series retain a noticeable level of multifractality, indicating that the non-Gaussian amplitude distribution remains a significant contributor. Moreover, the surrogated series exhibit a further decrease in multifractality, suggesting that nonlinear temporal correlations also play a significant role in generating the observed multifractal behavior in both indices across both periods.

Overall, the quote limits event does not fundamentally change the primary sources of multifractality in the returns of either the large-cap HS300 or small-cap CSI500 index. For both indices, the multifractal characteristics consistently stem from a combination of a non-trivial amplitude distribution and nonlinear temporal correlations, both before and after the implementation of the quote limits. These results suggest that the policy primarily influences the degree of multifractality rather than introducing or eliminating its fundamental sources.

TABLE 4 Change in sources of multifractality for stock index return.

Period		α_{min}	α_{max}	$\triangle\alpha$	$\triangle h$	*MDM*
Panel A: HS300						
Pre	Original	0.3119	0.9094	0.5975	0.3949	0.1975
	Shuffled	0.3493	0.8318	0.4824	0.3026	0.1513
	Surrogated	0.3964	0.7480	0.3516	0.2191	0.1095
Post	Original	0.0801	0.9896	0.9095	0.6699	0.3349
	Shuffled	0.1383	0.8135	0.6752	0.4631	0.2315
	Surrogated	0.3868	0.7270	0.3410	0.1999	0.0999
Panel B: CSI500						
Pre	Original	0.1880	1.0219	0.8339	0.5891	0.2946
	Shuffled	0.2609	0.8344	0.5734	0.3759	0.1879
	Surrogated	0.4753	0.7433	0.2679	0.1296	0.1047
Post	Original	0.1312	0.8824	0.7512	0.5523	0.2761
	Shuffled	0.2222	0.8484	0.6261	0.4389	0.2194
	Surrogated	0.3717	0.7922	0.4205	0.2495	0.1247

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after. "Original", "Shuffled", and "Surrogated" denote results for the original, shuffled, and surrogated spot and futures return series.

5 Effect of quote limits on cross-correlation between spot and futures market

5.1 Cross-correlation test

Following Kristoufek [52] and Ruan et al. [1], we first employ the DMA coefficient to quantify the linear cross-correlation between two time series at various time scales. Specifically, the DMA coefficient (ρ_{DMA}) is computed as follows:

$$\rho_{DMA} = \frac{F^2_{xy}(s)}{F^2_{xx}(s)F^2_{yy}(s)} \tag{15}$$

TABLE 5 Change in DMA coefficient.

s	4	8	16	32	64	128
Panel A: HS300 spot-futures						
Pre	0.9843	0.9903	0.9925	0.9937	0.9940	0.9942
Post	0.9827	0.9890	0.9923	0.9945	0.9946	0.9930
Panel B: CSI500 spot-futures						
Pre	0.9814	0.9889	0.9908	0.9911	0.9901	0.9868
Post	0.9824	0.9896	0.9929	0.9947	0.9953	0.9902

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after.

Accordingly, ρ_{DMA} values range from −1 to 1, with values close to 1 indicating strong positive linear correlation, values near −1 indicating strong negative linear correlation, and values around 0 suggesting weak or no linear correlation.

Table 5 presents the DMA coefficients between spot and futures returns for the HS300 and CSI500 indices before and after the implementation of quote limits. As shown, for the HS300 index, the DMA coefficients between spot and futures returns are consistently close to 1 before the quote limits event across all scales (s = 4–128). After the event, these coefficients remain high, with only minor fluctuations observed. Specifically, there is a slight decrease at shorter time scales (s = 4, 8) and a slight increase at longer time scales (s = 32, 64), with a small decrease at s = 128. Overall, the quote limits event appears to have a minimal impact on the strong positive linear cross-correlation between HS300 spot and futures returns.

Similarly, for the CSI500 index, the DMA coefficients between spot and futures returns exhibit strong positive linear correlation across all time scales before the event. Following the implementation of quote limits, the DMA coefficients remain high, showing a slight increase at shorter and intermediate scales (s = 4–64) and a slight increase followed by a decrease at the longest scale (s = 128). The magnitude of these changes is also small, indicating that the fundamental strong positive linear cross-correlation between CSI500 spot and futures returns was largely unaffected by the event.

Overall, the quote limits event does not significantly affect the strong positive linear relationship between the spot and futures returns for both the HS300 and CSI500.

5.2 Effect on the multifractality of spot-futures cross-correlations

Table 6 presents the generalized Hurst exponent $H(q)$, calculated using the MF-X-DMA method, for the cross-correlation between spot and futures returns for q ranging from −10 to 10. As shown, before the implementation of quote limits, the $H(q)$ for both HS300 and CSI500 spot-future cross-correlations vary with q, indicating the presence of multifractal characteristics. Specifically, the decreasing trend of $H(q)$ with increasing q suggests that large cross-correlation fluctuations scale differently from small ones.

TABLE 6 Change in the Hurst exponent behavior of spot-futures cross-correlations.

q	HS300 spot-futures		CSI500 spot-futures	
	Pre	Post	Pre	Post
−10	0.7684	0.8845	0.9393	0.8249
−9	0.7575	0.8704	0.9269	0.8148
−8	0.7445	0.8533	0.9118	0.8027
−7	0.7287	0.8325	0.8930	0.7882
−6	0.7094	0.8071	0.8693	0.7705
−5	0.6855	0.7768	0.8394	0.7485
−4	0.6564	0.7427	0.8023	0.7215
−3	0.6245	0.7074	0.7590	0.6895
−2	0.5975	0.6739	0.7126	0.6547
−1	0.5816	0.6434	0.6672	0.6208
0	0.5720	0.6124	0.6243	0.5880
1	0.5607	0.5689	0.5818	0.5477
2	0.5444	0.4955	0.5375	0.4890
3	0.5239	0.4071	0.4922	0.4238
4	0.5022	0.3366	0.4499	0.3710
5	0.4817	0.2881	0.4138	0.3328
6	0.4636	0.2545	0.3844	0.3047
7	0.4483	0.2301	0.3609	0.2833
8	0.4355	0.2117	0.3420	0.2665
9	0.4248	0.1972	0.3267	0.2529
10	0.4159	0.1855	0.3141	0.2417

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after.

After the quote limits event, the behavior of $H(q)$ diverges for the two respective spot-futures cross-correlations. For HS300 spot-future cross-correlations, $H(q)$ generally increases for negative q and decreases for large positive q, implying enhanced persistence in small cross-correlation fluctuations and stronger anti-persistence in large ones. Conversely, for CSI500 spot-futures cross-correlations, $H(q)$ generally decreases across the q range, suggesting a weakening of persistence for both small and large cross-correlation fluctuations.

To assess the impact of quote limits on the multifractality of spot-futures cross-correlation, we compute three indicators as in Table 3, i.e., the width of the multifractal spectrum ($\Delta\alpha$), the degree of the generalized Hurst exponent (Δh), and the Market Deficiency Measure (MDM). Accordingly, higher values of these

TABLE 7 Change in multifractal spectra characteristics for spot-futures cross-correlations.

Period	α_{min}	α_{max}	$\triangle\alpha$	$\triangle h$	*MDM*
Panel A: HS300 spot-futures					
Pre	0.3266	0.8774	0.5508	0.3525	0.1762
Post	0.0685	1.0250	0.9565	0.6990	0.3495
Panel B: CSI500 spot-futures					
Pre	0.1882	1.0633	0.8751	0.6252	0.3126
Post	0.1295	0.9264	0.7969	0.5833	0.2916

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after.

indicators signify stronger multifractality, implying more complex and heterogeneous co-movement across different fluctuation magnitudes.

Table 7 reports the results for HS300 and CSI500 spot-futures cross-correlations before ajnd after the event. As shown, post-event values for $\Delta\alpha$, Δh, and *MDM* for the HS300 spot-futures cross-correlation all exhibit a substantial increase compared to pre-event levels. This increase in all three measures suggests that the quote limits policy leads to a higher degree of multifractality in the relationship between HS300 spot and futures returns, indicating a more complex co-movement across varying fluctuation magnitudes. In contrast, for the CSI500 spot-futures cross-correlation, the post-event values for $\Delta\alpha$, Δh, and *MDM* all show a decrease from their pre-event levels. This consistent decrease in multifractal indicators suggests that the implementation of quote limits results in a reduced degree of multifractality in the co-movement between CSI500 spot and futures returns, potentially indicating a simplification of their relationship. The multifractal spectrum depicted in Figure 2 also shows a similar pattern.

In sum, the implementation of quote limits has a heterogeneous influence on the dynamics between the derivative and underlying markets for these two key Chinese equity indices. It increases the complexity of the co-movement of HS300 spot-futures returns, indicating less efficient cross-market information transmission. In contrast, it decreases the complexity of the co-movement of CSI500 spot-futures returns, leading to a more efficient cross-market information transmission. These results can also help explain our findings regarding the divergent impact of quote limits on individual market efficiency, i.e., increased complexity for HS300 spot-futures cross-correlation aligns with decreased efficiency of the HS300 index, and *vice versa* for CSI500.

5.3 Change in multifractal sources of spot-futures cross-correlations

Using the same methodology outlined in Section 4.3, we then examine whether the sources of multifractality in spot-futures cross-correlations are affected by the implementation of quote limits. Table 8 presents the multifractal measures for the original, shuffled, and surrogated series before and after the event.

For the HS300 spot-futures cross-correlation, the multifractal sources exhibit a notable change. Before the event, both the shuffled and surrogated series show reduced multifractality compared to the original, indicating contributions from both the non-Gaussian amplitude distribution and nonlinear temporal dependencies. After the event, while the original series display stronger multifractality, the surrogated series show a substantial increase in multifractality. This suggests that linear temporal correlations become a more significant driver of multifractality in the HS300 spot-future cross-correlation following the implementation of quote limits. The shuffled series, in contrast, show a much lower level of multifractality, implying a diminished role of the amplitude distribution alone.

The multifractal sources for the CSI500 spot-futures cross-correlation exhibit a different pattern. Before the event, similar to the HS300 spot-futures returns, both shuffled and surrogated series showed reduced multifractality. After the event, the shuffled series still exhibited a considerable level of multifractality, suggesting a persistent role of the amplitude distribution. The surrogated series showed a modest increase in multifractality compared to its pre-event level, indicating a potentially slightly enhanced role of linear temporal correlations.

Comparing the two indices, the quote limits event appears to have a distinct impact on the sources of multifractality in their spot-futures cross-correlations. For HS300, the increased multifractality seems to be primarily driven by linear temporal correlations. For CSI500, while there is a slight increase in the contribution from linear temporal correlations, the amplitude distribution remains a relevant source of multifractality.

5.4 Change in nonlinear granger causality

Nonlinear Granger causality tests whether past values of one return series contain unique nonlinear information that improves the prediction of another series' current values [28, 37, 53, 54]. In financial contexts, this implies that certain past price movements or trading activities in one market (e.g., futures) can provide nonlinear predictive power for another related market (e.g., spot), reflecting potentially asymmetric or complex information transmission and market efficiency dynamics [1, 15, 24, 55].

Table 9 presents results of the Nonlinear Granger causality test. For the HS300 spot and futures return series, significant nonlinear Granger causality runs unidirectionally from futures to spot returns at lag 1 before the implementation of quote limits. After the event, this unidirectional causality persists at lag 1 and becomes significant at lags 2, 3, and 4 (at the 10% level). There is no evidence of nonlinear causality from spot to futures in either period. Consequently, the quote limits event appears to strengthen and broaden the nonlinear influence of HS300 futures returns on spot returns across multiple lags. This may be due to the decreased market efficiency of the HS300 spot market, thereby strengthening the informational leadership of the futures market over the spot market.

Regarding the CSI500 index, before the event, there exists significant nonlinear Granger causality from futures to spot

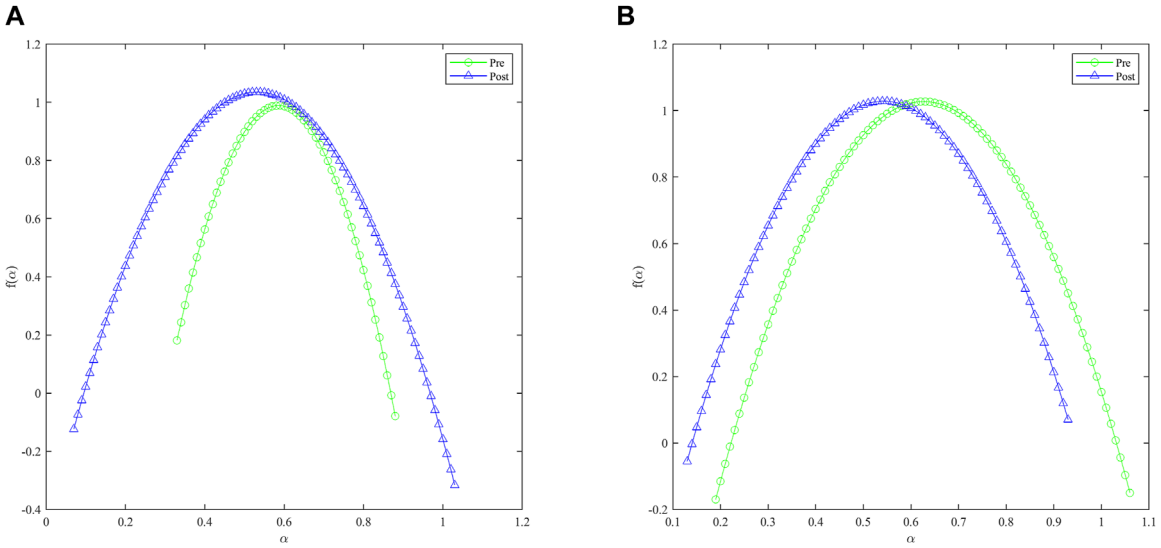


FIGURE 2
Change in Multifractal spectra for spot-futures cross-correlations. **(A)**: HS300 spot-futures. **(B)**: CSI500 spot-futures.

returns at lags 1 and 2. Post-event, this unidirectional causality not only remains significant but also strengthens across lags 1 to 4. Notably, a bidirectional nonlinear Granger causality emerges at lags 1 and 2, with significant causality running from spot to futures as well. Therefore, the quote limits event fundamentally altered the nonlinear causal relationship between CSI500 spot and futures returns, establishing a more integrated and sophisticated bidirectional information flow at shorter lags. This finding aligns with the results of increased market efficiency and reduced complexity for the CSI500 spot market, implying a more mature relationship between the derivative and its underlying.

Overall, the quote limits policy also has a distinct impact on their nonlinear causal relationships. Specifically, it primarily enhances the existing unidirectional causality from futures to spot for HS300, whereas it leads to a more significant change by establishing a bidirectional nonlinear causality at shorter lags, alongside the strengthening of the futures-to-spot causality for CSI500.

TABLE 8 Change in sources of multifractality for spot-futures cross-correlations.

Period		α_{min}	α_{max}	$\triangle\alpha$	$\triangle h$	*MDM*
Panel A: HS300 spot-futures						
Pre	Original	0.3266	0.8774	0.5508	0.3525	0.1762
	Shuffled	0.3385	0.7266	0.3881	0.2275	0.1137
	Surrogated	0.4056	0.6574	0.2518	0.1277	0.0639
Post	Original	0.0685	1.0250	0.9565	0.6990	0.3495
	Shuffled	0.5347	0.7699	0.2352	0.1175	0.1297
	Surrogated	0.3189	0.8635	0.5446	0.3484	0.1742
Panel B: CSI500 spot-futures						
Pre	Original	0.1882	1.0633	0.8751	0.6252	0.3126
	Shuffled	0.3628	0.7727	0.4099	0.2581	0.1291
	Surrogated	0.4594	0.8233	0.3640	0.1996	0.1119
Post	Original	0.1295	0.9264	0.7969	0.5833	0.2916
	Shuffled	0.3005	1.0361	0.7356	0.5168	0.2584
	Surrogated	0.2948	0.7585	0.4638	0.2849	0.1425

Note: "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after. "Original", "Shuffled", and "Surrogated" denote results for the original, shuffled, and surrogated spot and futures return series.

6 Conclusion and discussion

This study investigated the impact of the quote limits policy on market efficiency and the complex dynamics between spot and futures markets of the HS300 and CSI500 indices. Using the MF-X-DMA approach and nonlinear Granger causality tests, our findings reveal a significantly different impact on these two key Chinese equity benchmarks.

Specifically, the implementation of quote limits decreases the market efficiency of the large-cap HS300 index, evidenced by a marked increase in multifractal characteristics in HS300 index returns. For the HS300 spot-futures relationship, while linear correlations remain robust, the complexity of co-movement increases, and nonlinear Granger causality from futures to spot is strengthened and broadened, suggesting a more dominant and potentially asymmetric information flow from the derivatives

TABLE 9 Change in Non-linear Granger causality between spot and futures returns.

		HS300		CSI500	
Lag	Period	Spot $\nrightarrow$ futures	Futures $\nrightarrow$ spot	Spot $\nrightarrow$ futures	Futures $\nrightarrow$ spot
1	Pre	−0.4686	1.6464**	0.5315	1.6925**
	Post	0.8401	1.9353**	1.7958**	2.0464**
2	Pre	0.8371	0.7983	0.3669	1.3624*
	Post	0.8650	1.3217*	1.9196**	2.316**
3	Pre	0.1554	0.2776	0.3728	1.2698
	Post	−0.1009	1.3105*	0.8162	1.9696**
4	Pre	0.1385	0.9918	0.3141	0.8862
	Post	0.3871	1.5064*	0.6921	1.8172**

Note: "spot $\nrightarrow$ futures" denotes the null hypothesis that spot returns do not nonlinearly Granger cause futures returns, while "futures $\nrightarrow$ spot" denotes the null hypothesis that futures returns do not nonlinearly Granger cause spot returns. "Pre" refers to the sub-period before the implementation of quote limits, and "Post" refers to the sub-period after. "*," "**," and "***" suggest rejection of the null hypothesis at the 10%, 5%, and 1% significance levels, respectively.

market. Notably, for HS300, the increased multifractality in spot-futures cross-correlations appears to be primarily driven by linear temporal correlations.

In contrast, the policy enhances the market efficiency of the mid-to-small cap CSI500 index, reflected in a reduction in multifractal characteristics in the CSI500 index returns. More importantly, the cross-market analysis shows a simplification of co-movement complexity and the emergence of bidirectional nonlinear Granger causality between spot and futures returns. This indicates the development of a more interactive and bidirectional information flow, where both markets dynamically influence each other. For CSI500, the sources of multifractality in spot-futures cross-correlations show a slight increase in the contribution from linear temporal correlations, though the amplitude distribution remains a relevant source.

From a policy perspective, our findings highlight that a "one-size-fits-all" approach to market regulation may yield unintended consequences. While the quote limits policy achieves its goal of stabilizing certain market segments or curtailing excessive speculation (as seen in CSI500s increased efficiency), it simultaneously introduces new complexities or reduced efficiency in others (as observed in HS300). This underscores the critical need for policymakers to adopt tailored regulatory designs that account for the unique characteristics of different market segments and to continuously assess the actual impact of regulations. For market participants, our results imply the necessity of developing segment-specific trading strategies and adaptive risk management frameworks, given the distinct alterations in market dynamics for the HS300 *versus* the CSI500.

A possible concern is that changes in multifractal characteristics can be influenced by various complex factors and may not represent the sole or absolute measure of market efficiency. Furthermore, while this study reveals the impact of the quote limits system at the index level, future research could further explore the system's effects on the multifractal characteristics of individual stocks. This would allow for more granular insights into market dynamics and help prevent the loss of detailed information that can occur through index-level averaging.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Requests to access these datasets should be directed to Linsen Yin, yls@lixin.edu.cn.

Author contributions

DL: Funding acquisition, Writing – original draft, Formal Analysis, Software, Writing – review and editing, Visualization, Data curation, Validation, Conceptualization, Methodology. YZ: Writing – review and editing, Conceptualization, Writing – original draft, Validation, Data curation, Methodology, Formal Analysis. JZ: Validation, Data curation, Writing – review and editing, Methodology. LY: Validation, Project administration, Conceptualization, Writing – original draft, Resources, Formal Analysis, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors wish to express their gratitude for the support provided by

the Shuguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission (No. 22SG55), the Humanities and Social Science Fund of Ministry of Education of the People's Republic of China (No. 21YJC790082), and the National Natural Science Foundation of China (No. 72103139).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Ruan Q, Hu S, Zhang J, Chu X, Lv D. Relaxing daily price limits and stock market cross-correlation: evidence from MF-X-DMA analysis. *Fluctuation Noise Lett* (2023) 22:2350042. doi:10.1142/s0219477523500426

2. Wang X, Kim MH, Suardi S. Herding and China's market-wide circuit breaker. *J Banking and Finance* (2022) 141:106533. doi:10.1016/j.jbankfin.2022.106533

3. Switzer LN, Yue H. Effects of the short sale circuit breaker on the stock market. *J Int Financial Manage and Account* (2019) 30:250–74. doi:10.1111/jifm.12105

4. Deb SS, Kalev PS, Marisetty VB. Are price limits really bad for equity markets? *J Banking and Finance* (2010) 34:2462–71. doi:10.1016/j.jbankfin.2010.04.001

5. Liang H, Sun Y, Xu C, Xiong W, Cai W. Unleashing stock volatility and its implications for stock crash risk: evidence from China's price limit policies. *Res Int Business Finance* (2024) 71:102455. doi:10.1016/j.ribaf.2024.102455

6. Heflin F, Shaw KW. Trade size and informed trading: which trades are "big". *J Finan Res* (2005) 28:133–63. doi:10.1111/j.1475-6803.2005.00118.x

7. Bogousslavsky V, Fos V, Muravyev D. Informed trading intensity. *J Finance* (2024) 79:903–48. doi:10.1111/jofi.13320

8. Chan K, Fong W-M. Trade size, order imbalance, and the volatility–volume relation. *J Finan Econ* (2000) 57:247–73. doi:10.1016/S0304-405X(00)00057-X

9. Lee J, Yi CH. Trade size and information-motivated trading in the options and stock markets. *J Finan Quant Anal* (2001) 36:485–501. doi:10.2307/2676221

10. Zhao R, Cui Y, Liu X. Tick size and market quality using an agent-based multiple-order-book model. *Front Phys* (2020) 8:8–2020. doi:10.3389/fphy.2020.00135

11. Lu X, Tian J, Zhou Y, Li Z. Multifractal detrended fluctuation analysis of the Chinese stock index futures market. *Physica A: Stat Mech its Appl* (2013) 392:1452–8. doi:10.1016/j.physa.2012.11.037

12. Zhou W-X. Finite-size effect and the components of multifractality in financial volatility. *Chaos, Solitons and Fractals* (2012) 45:147–55. doi:10.1016/j.chaos.2011.11.004

13. Lee M-J, Choi S-Y. Comparing market efficiency in developed, emerging, and frontier equity markets: a multifractal detrended fluctuation analysis. *Fractal and Fractional* (2023) 7:478. doi:10.3390/fractalfract7060478

14. Jiang Z-Q, Xie W-J, Zhou W-X, Sornette D. Multifractal analysis of financial markets: a review. *Rep Prog Phys* (2019) 82:125901. doi:10.1088/1361-6633/ab42fb

15. Kojić M, Mitić P, Antunes De Araujo FH. Multiscale nonlinearity in energy commodities and bond markets: MFDCCA via MODWT and granger causality. *Fractals* (2025) 33(0):2550030. doi:10.1142/s0218348x25500306

16. Acikgoz T. The multifractal nature of cross-correlations between emerging market equities and financial assets: an econophysics perspective. *Comput. Econ.* Netherlands: Springer (2025). doi:10.1007/s10614-025-10991-4

17. Zhao R, Dai P-FA. A multifractal cross-correlation analysis of economic policy uncertainty: evidence from China and US. *Fluctuation Noise Lett* (2021) 20:2150041. doi:10.1142/s0219477521500413

18. Ghosh I, Chaudhuri TD, Sarkar S, Mukhopadhyay S, Roy A. Macroeconomic shocks, market uncertainty and speculative bubbles: a decomposition-based predictive model of Indian stock markets. *China Finance Rev Int* (2025) 15:166–201. doi:10.1108/CFRI-09-2023-0237

19. Zhang L, Hua L. Market predictability before the closing bell rings. *Risks* (2024) 12:180. doi:10.3390/risks12110180

20. Gu D, Huang J. Multifractal detrended cross-correlation analysis of high-frequency stock series based on ensemble empirical mode decomposition. *Fractals* (2020) 28:2050035. doi:10.1142/s0218348x20500358

21. Zhou W, Huang J, Wang M. Multifractal characteristics and information flow analysis of stock markets based on multifractal detrended cross-correlation analysis and transfer entropy. *Fractal and Fractional* (2025) 9:14. doi:10.3390/fractalfract9010014

22. Ye S, Dai P-F, Nguyen HT, Huynh NQA. Is the cross-correlation of EU carbon market price with policy uncertainty really being? A multiscale multifractal perspective. *J Environ Manage* (2021) 298:113490. doi:10.1016/j.jenvman.2021.113490

23. Hou L, Pan Y. Evaluating the connectedness of commodity future markets via the cross-correlation network. *Front Phys* (2022) 10:10–2022. doi:10.3389/fphy.2022.1017009

24. Zhang P, Yin S, Sha Y. Global systemic risk dynamic network connectedness during the COVID-19: evidence from nonlinear Granger causality. *J Int Finan Markets, Inst Money* (2023) 85:101783. doi:10.1016/j.intfin.2023.101783

25. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of DNA nucleotides. *Phys Rev E* (1994) 49:1685–9. doi:10.1103/PhysRevE.49.1685

26. Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Stat Mech its Appl* (2002) 316:87–114. doi:10.1016/S0378-4371(02)01383-3

27. Podobnik B, Stanley HE. Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. *Phys Rev Lett* (2008) 100:084102. doi:10.1103/PhysRevLett.100.084102

28. Zhou W-X. Multifractal detrended cross-correlation analysis for two nonstationary signals. *Phys Rev E* (2008) 77:066211. doi:10.1103/PhysRevE.77.066211

29. Jiang Z-Q, Zhou W-X. Multifractal detrending moving-average cross-correlation analysis. *Phys Rev E* (2011) 84:016106. doi:10.1103/PhysRevE.84.016106

30. Li J, Lu X, Qu L. Effectiveness of the RMB exchange rate regime reform: a new perspective from MF-DMA and MF-X-DMA. *Physica A: Stat Mech its Appl* (2019) 531:121535. doi:10.1016/j.physa.2019.121535

31. Mikhaylov A, Dinçer H, Yüksel S. Analysis of financial development and open innovation oriented fintech potential for emerging economies using an integrated decision-making approach of MF-X-DMA and golden cut bipolar q-ROFSs. *Financial Innovation* (2023) 9:4. doi:10.1186/s40854-022-00399-6

32. Baek E, Brock W. *A general test for nonlinear Granger causality: bivariate model.* Iowa State University and University of Wisconsin at Madison Working Paper (1992).

33. Diks C, Wolski M. Nonlinear Granger causality: guidelines for multivariate analysis. *J Appl Econom* (2016) 31:1333–51. doi:10.1002/jae.2495

34. Zhou Y, Lu B, Lv D, Ruan Q. The informativeness of options-trading activities: non-linear analysis based on MF-DCCA and Granger test. *Physica A: Stat Mech its Appl* (2019) 534:122269. doi:10.1016/j.physa.2019.122269

35. Vogl M, Kojić M. Green cryptocurrencies versus sustainable investments dynamics: exploration of multifractal multiscale analysis, multifractal detrended cross-correlations and nonlinear Granger causality. *Physica A: Stat Mech its Appl* (2024) 653:130085. doi:10.1016/j.physa.2024.130085

36. Tao J, Yingying W, Jingyi Z. The performance of China's stock market price limits: noise mitigator or noise maker? *China Finance Rev Int* (2017) 7:85–97. doi:10.1108/CFRI-07-2016-0096

37. Cao G, Xu W. Nonlinear structure analysis of carbon and energy markets with MFDCCA based on maximum overlap wavelet transform. *Physica A: Stat Mech its Appl* (2016) 444:505–23. doi:10.1016/j.physa.2015.10.070

38. Wang F, Ye X, Chen H, Wu C. A portfolio strategy of stock market based on mean-MF-X-DMA model. *Chaos, Solitons and Fractals* (2021) 143:110645. doi:10.1016/j.chaos.2020.110645

39. Wang C, Wang C, Wu W. Customers' stock price crash risk and suppliers' investment inefficiency: evidence from China. *Account and Finance* (2023) 63:5069–92. doi:10.1111/acfi.13143

40. Titman S, Wei C, Zhao B. Corporate actions and the manipulation of retail investors in China: an analysis of stock splits. *J Finan Econ* (2022) 145:762–87. doi:10.1016/j.jfineco.2021.09.018

41. Carpenter JN, Lu F, Whitelaw RF. The real value of China's stock market. *J Finan Econ* (2021) 139:679–96. doi:10.1016/j.jfineco.2020.08.012

42. Boungou W, Yatié A. The impact of the Ukraine–Russia war on world stock market returns. *Econ Lett* (2022) 215:110516. doi:10.1016/j.econlet.2022.110516

43. Vidal-Tomás D. The new crypto niche: NFTs, play-to-earn, and metaverse tokens. *Finance Res Lett* (2022) 47:102742. doi:10.1016/j.frl.2022.102742

44. Grobys K, Junttila J. Speculation and lottery-like demand in cryptocurrency markets. *J Int Finan Markets, Inst Money* (2021) 71:101289. doi:10.1016/j.intfin.2021.101289

45. Zunino L, Tabak BM, Figliola A, Pérez DG, Garavaglia M, Rosso OA. A multifractal approach for stock market inefficiency. *Physica A: Stat Mech its Appl* (2008) 387:6558–66. doi:10.1016/j.physa.2008.08.028

46. Fernandes LHS, de Araújo FHA, Silva IEM. The (in)efficiency of NYMEX energy futures: a multifractal analysis. *Physica A: Stat Mech its Appl* (2020) 556:124783. doi:10.1016/j.physa.2020.124783

47. Castro AL, Marcato ALM, Aguiar EPD. Multifractal analysis of the Brazilian electricity market. *IEEE Access* (2023) 11:98939–57. doi:10.1109/ACCESS.2023.3313099

48. Ock SE, Lee M, Song JW. Multifractal analysis of the impact of fuel cell introduction in the Korean electricity market. *Fractal and Fractional* (2024) 8:573. doi:10.3390/fractalfract8100573

49. Kwapień J, Blasiak P, Drożdż S, Oświęcimka P. Genuine multifractality in time series is due to temporal correlations. *Phys Rev E* (2023) 107:034139. doi:10.1103/PhysRevE.107.034139

50. Rak R, Grech D. Quantitative approach to multifractality induced by correlations and broad distribution of data. *Physica A: Stat Mech its Appl* (2018) 508:48–66. doi:10.1016/j.physa.2018.05.059

51. Drożdż S, Kwapień J, Oświecimka P, Rak R. Quantitative features of multifractal subtleties in time series. *Europhysics Lett* (2009) 88:60003. doi:10.1209/0295-5075/88/60003

52. Kristoufek L. Detrending moving-average cross-correlation coefficient: measuring cross-correlations between non-stationary series. *Physica A: Stat Mech its Appl* (2014) 406:169–75. doi:10.1016/j.physa.2014.03.015

53. Shojaie A, Fox EB. Granger causality: a review and recent advances. *Annu Rev Stat Its Appl* (2022) 9:289–319. doi:10.1146/annurev-statistics-040120-010930

54. Diks C, Panchenko V. A new statistic and practical guidelines for nonparametric Granger causality testing. *J Econ Dynam Control* (2006) 30:1647–69. doi:10.1016/j.jedc.2005.08.008

55. Li S, Tu D, Zeng Y, Gong C, Yuan D. Does geopolitical risk matter in crude oil and stock markets? Evidence from disaggregated data. *Energy Econ* (2022) 113:106191. doi:10.1016/j.eneco.2022.106191

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 30 July 2025
DOI 10.3389/fphy.2025.1610832

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Jionghao Huang,
Peking University, China
Min Pang,
Southwest Jiaotong University, China

*CORRESPONDENCE
Qing Guan,
guanqing35@126.com

RECEIVED 13 April 2025
ACCEPTED 11 July 2025
PUBLISHED 30 July 2025

CITATION
Wang Y and Guan Q (2025) Studies on innovative energy startups' topological roles and their correlation with success: based on temporal networks.
Front. Phys. 13:1610832.
doi: 10.3389/fphy.2025.1610832

COPYRIGHT
© 2025 Wang and Guan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Studies on innovative energy startups' topological roles and their correlation with success: based on temporal networks

Ying Wang[1] and Qing Guan[2]*

[1]School of Economics and Management, China University of Geosciences, Beijing, China, [2]School of Artificial Intelligence, China University of Geosciences, Beijing, China

Introduction: Innovative energy startups are expediting the energy transition through the adoption of emerging technologies, including blockchain, fintech, artificial intelligence, and crowdfunding. However, existing research primarily focuses on technological capabilities at the startup level and macro-level national applications to explore the current state of innovative energy adoption. In contrast, limited attention has been paid to analyzing the role attributes of innovative energy startups and their correlations with potential success, which are critical for understanding their development trajectories within the energy market.

Methods: This study develops a temporal investment information network for global energy startups, drawing on data from energy enterprises worldwide between 2005 and 2024. The research examines the role attributes of startups and explores the temporal topological characteristics of the network. We propose a success evaluation model based on the features of successful startups to assess the potential of innovative energy startups.

Results and Discussion: The findings indicate that, despite their relatively small market share, innovative energy startups exert significant influence. Notably, successful startups typically exhibit higher betweenness centrality and lower closeness centrality. Moreover, factors such as network degree, centrality, and government administrative capacity play crucial roles in determining the success of innovative energy startups. In the evaluation model constructed using these factors, network structural characteristics contribute the most, achieving an evaluation accuracy of 0.984. This study provides valuable insights for policymakers evaluating innovative energy development trends and for investors assessing the potential of startups.

KEYWORDS

innovative energy startups, temporal networks, investment information, success relevance, development direction

1 Introduction

Innovative energy has become a forefront issue on the global economic and environmental agenda due to its ability to ensure economic competitiveness [1, 2], improve energy efficiency [3], reduce environmental pollution, and mitigate climate

change [4]. Among them, innovative energy startups serve as a new perspective for studying energy market development by analyzing their market return effects and connectivity [5], risk spillover effects [6], extreme uncertainty connectedness [7], and environmental sustainability and its development [8, 9] in comparison to other types of startups. The individual roles and topological characteristics of startups can reflect system structure and interaction patterns [10], thereby revealing interconnections, dependencies, and interactions among startups. However, existing literature primarily examines innovative energy startups from a national perspective [11] and a technological perspective [12–14], neglecting the individual roles and interrelationships of these startups. Therefore, depicting the roles and positions of innovative energy startups within the business ecosystem is of significant importance.

One of the goals of startup development is business success. Existing research analyzes whether a startup can succeed from two main perspectives. First, from a micro-level perspective focusing on individual startups, studies have identified success factors such as innovative business models and collaborations, while failure factors include strong incumbent industries and unfavorable investor sentiment [15, 16]. Second, from a macro-level systemic perspective, research explores investment information on innovative energy startups [17, 18], characterizes static network capital through the final network structure [19], examines communication patterns among startups [20, 21], and uncovers potential dependencies, such as knowledge linkages [22], to assess the likelihood of future success [23–25]. In the energy market, traditional social network analysis (SNA) is suitable for examining stable structural patterns, such as the relationship between energy and capabilities [26] and the fixed spatial distribution of energy consumption [27]. However, due to the rapid environmental changes and increasing competitive pressures driven by technological innovations [28], the success of innovative energy startups increasingly depends on the ability to capture and analyze the dynamic characteristics of the energy market. Temporal networks have demonstrated advantages in capturing dynamic network changes, whether in studying startup growth [29] or optimizing stock market investment portfolios [30]. Therefore, by depicting the dynamic topological characteristics of innovative energy startups through the network structure of temporal networks and leveraging network features to analyze successful energy startups and their connections with innovative energy startups, it is possible to explore the future development directions of innovative energy startups.

To address the existing gaps in understanding the market dynamics and success mechanisms of energy startups, this study introduces an innovative framework that integrates both static and temporal network perspectives. Drawing on a comprehensive dataset of 26,188 investment records, we construct global static and temporal investment networks encompassing 3,592 energy startups connected through shared investors. This network approach reveals both the market structure of innovative energy startups and their evolving connections with successful startups over time, using key metrics such as betweenness and closeness centrality. Building on structural positioning in the networks, we evaluate the role attributes of startups and examine the correlation between network properties and startup success. Furthermore, we develop a novel correlation analysis model to assess success potential in the energy startup. Finally, we apply this model to explore the future development paths of innovative energy startups, providing new insights into the dynamic evolution of the energy startup landscape.

The research work of this paper is as follows. In Section 2, we briefly review the relevant literature and emphasize the necessity of incorporating dynamic network attributes into the analysis of the success correlation of innovative energy startups. Section 3 presents the data. In Section 4, we outline the construction of the network model and the evaluation model, along with their analytical methods. Section 5 discusses the results. Finally, in Section 6, we provide a summary and propose directions for future research.

2 Literature review

The role of innovative energy in the global energy market is increasingly recognized [31, 32]. As one of the key subjects in innovative energy research, startups are studied across various dimensions, including technology types [33, 34], real-world applications [35], investment [36], regional distribution [37], and policy design [38]. Among these factors, the role of government in promoting innovative energy startups is crucial. Governments support these startups by implementing proactive regulatory policies, funding private enterprises, offering innovation awards, and relaxing market constraints [39, 40]. This support plays a significant role in fostering their growth [41]. However, despite the supportive role of government in advancing energy technologies, a significant commercialization gap in energy innovation remains [40, [42]]. In practice, the successful deployment and diffusion of energy innovation rely heavily on the implementation capabilities of startups, and the effectiveness of policy measures often varies depending on firm-specific characteristics. Yet, existing research tends to overlook the individual roles and features of innovative energy startups, which are crucial for driving the development of energy innovation.

Investment activities play a pivotal role in the growth trajectory of startups. However, the absence of historical benchmarks, the volatility of entrepreneurial ecosystems, and the rapid movement of capital continue to pose significant challenges to successful investment outcomes [43, 44]. Existing literature primarily focuses on assessing entrepreneurial success through financial and operational indicators [82], with funding capacity often used as a proxy—reflected in outcomes such as initial public offerings [45], financing rounds [46], and acquisitions [47, 48]. From the perspective of founding teams, factors such as prior experience, educational background, skills, and even personality traits have been shown to influence startup performance [49, 50]. From the investor's standpoint, although startup investment entails high risk, investors often play a central role in critical strategic decisions [51, 52]. However, despite the pivotal role of investment in startup success, existing research has paid relatively limited attention to the mechanisms through which investment-related information functions, particularly in the context of innovative energy startups [53, 54]. The impact of investment information on the developmental trajectories of such startups remains an underexplored area worthy of further investigation.

There are complex interactions among innovative energy startups. Through network analysis based on investment information, we can characterize the role attributes and topological features of innovative energy startups from a macro perspective. Role attributes refer to the relational assets generated by an individual's interactions within a complex external network [55]. Establishing network connections with external resources plays a crucial role in facilitating information exchange [56], which helps explore the development of individuals within the network [19]. In practice, the network effects generated by role attributes facilitate information transmission between indirectly connected individuals through intermediaries [19]. For example, although Startup A and Startup C may not have a direct connection, technology [57], knowledge [58], and information [22] can be exchanged through an intermediary, Startup B. Increasingly, network-based technology diffusion models have explored how information spreads across different network configurations, highlighting the importance of role attributes in information acquisition [59–62]. The theory of preferential attachment [63] further explains this mechanism by suggesting that firms with stronger role attributes are more likely to attract new connections. As a result, these firms become increasingly central within the network structure. Moreover, some studies explore energy startups through the topological features of networks. For example, [19] constructed a global diffusion network for climate change mitigation technologies using patent data, company relationships, and business scale in their study of energy startups. By calculating network attributes such as degree, closeness centrality, and betweenness centrality, they found significant network effects in link quantity and structural positioning. Similarly, [64] analyzed the Chinese energy industry and energy startup market using network metrics such as degree, closeness centrality, and betweenness centrality, uncovering spillover effects in the stock market and the role of carbon markets in network benefits. However, these studies analyze energy startups solely based on the final static network structure and overlook the impact of market dynamics such as changes in investment information on the network role attributes and topological features of individual startups.

Temporal networks enable the dynamic characterization of network roles by continuously tracking key nodes [65, 66] and their associated nodes over time [67]. Therefore, this study introduces an analytical framework that integrates both static and temporal network perspectives, using investment information to construct a dynamically evolving investment network that captures the individual attributes of innovative energy startups across different time periods. By analyzing the topological features of successful startups, we further identify key structural characteristics that influence the growth of energy startups. This approach not only overcomes the limitations of traditional static analyses by capturing the impact of market dynamics on startup roles, but also provides a novel tool for exploring the developmental trajectories of innovative energy startups.

3 Data

This study constructs an investment information network for energy startups using global corporate data sourced from the Crunchbase database (https://www.crunchbase.com/), covering the period from January 2005 to July 2024. The dataset includes critical information such as announcement dates, investor names, funding status, organization industries, country-level geographic data, and administrative capacity data of energy startups.

We categorize energy market startups into three types: traditional energy startups, renewable energy startups, and innovative energy startups. Innovative energy startups refer to those that integrate technologies such as blockchain, fintech, artificial intelligence, and crowdfunding into their organizational structures and industrial domains to drive the transformation of the energy sector [68–70]. Renewable energy startups are defined as those whose organizational industries include the renewable energy sector, while all other startups are classified as traditional energy startups by default. Furthermore, we classify energy market startups based on their success status. Unlike previous studies that measure success using economic indicators such as corporate performance [71], we define a successful energy startup as one that has achieved either an initial public offering (IPO) or has been acquired through a merger or acquisition (M&A). This binary classification is widely used in innovation and entrepreneurship literature as a practical proxy for success, due to the clear and observable nature of these outcomes [72]. Ultimately, our dataset comprises 295 innovative energy startups (3.3%), 3,488 renewable energy startups (39.2%), and 5,117 traditional energy startups (57.5%). Among them, only 1,507 startups (16.9%) are classified as successful, highlighting the considerable challenges startups face in achieving success.

Government administrative capacity is measured by government effectiveness and regulatory quality. This variable reflects the government's ability to regulate the local energy startup market and the ease with which investors can conduct business in the country [19]. Data on national administrative capacity are sourced from the Worldwide Governance Indicators provided by the World Bank (https://www.worldbank.org/). Government effectiveness captures the credibility of the government's commitment to policies and the quality of their implementation, while regulatory quality reflects the government's ability to formulate laws and regulations that enable and promote private sector development [73]. We take the average of the estimated governance values for each country from 1996 to 2022 as a measure of its government effectiveness and regulatory quality.

4 Methods

The methodological framework of this study consists of three components. In Section 4.1, we construct a static network of global energy startup investment information and employ network statistical analysis to examine the overall market characteristics of energy startups. In Section 4.2, we build a temporal network of investment information between startups to analyze the dynamic market characteristics of innovative and successful energy startups. In Section 4.3, we develop an evaluation model based on the network variables from Section 4.2, exploring the relationship between these variables and startup success while assessing the accuracy of the evaluation model.

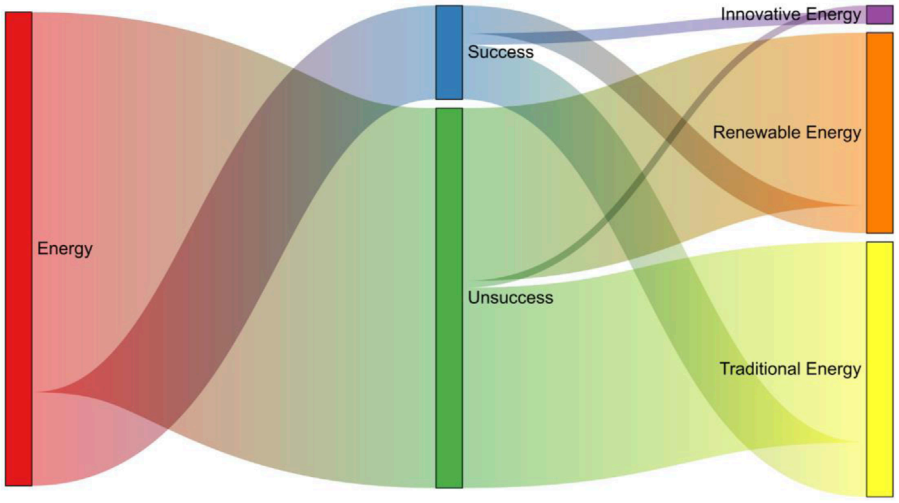


FIGURE 1
Energy startups market distribution.

TABLE 1 Scale of energy startups and distribution of government capacity in the top six countries.

Country	Innovative energy	Renewable energy	Traditional energy	Success	Government effectiveness	Regulatory quality
USA	56	652	918	343	1.521	1.474
Canada	11	117	207	147	1.769	1.646
UK	11	110	139	30	1.601	1.722
Germany	11	61	65	22	1.553	1.572
Spain	1	36	102	138	1.175	1.066
China	1	25	78	14	1.010	1.014

4.1 Static investment information network analysis

We construct a static investment information network for energy startups, where startups are represented as nodes, and edges indicate the presence of shared investors between startups. In this section, we treat the network as an outcome rather than a process and analyze the global distribution of investment resources in the energy market by characterizing the static features of startups.

We construct static network where startups are represented as nodes, and edges indicate the presence of shared investors between startups. Startups without shared investors are removed, resulting in a final network comprising 3,592 energy startup nodes and 26,188 edges. Among these, there are 138 innovative energy startups, 1,520 renewable energy startups, 1,936 traditional energy startups, and 711 successful startups, as shown in Figure 1. Ranking countries by the number of energy startups, we find that the top six countries account for 70% of the market's startups, with over 90% of successful startups located in these countries. Detailed information is presented in Table 1.

Edges reflect the interactions between startups through investment information. We measure edges using investment data between startups, including in-degree, out-degree, weighted in-degree, and weighted out-degree. Specifically, in-degree represents the number of investments received from investors of other startups, indicating a startup's attractiveness to investors. Out-degree represents the number of investments made by a startup's investors in other startups, reflecting the extent of capital spillover. The weight of an edge represents the frequency of investment transactions, capturing the intensity of investment activity. Mathematically, following [74], we define a network $G(V,E)$, where $V=\{v_1,v_2,\ldots,v_n\}$ denotes the set of nodes, v_n representing different startups, and E represents the set of edges, which can be expressed in matrix form, as shown in Equation 1:

$$E=\begin{bmatrix} e_{11} & \cdots & e_{1n} \\ \vdots & \ddots & \vdots \\ e_{m1} & \cdots & e_{mn} \end{bmatrix} \tag{1}$$

where e_{mn} is the link from startups v_m to v_n,which can be defined as shown in Equation 2:

$$e_{mn} = \begin{cases} 0 \; if \; there \; is \; no \; common \; investor \; between \; v_m \; and \; v_n \\ w_{mn} \; if \; investor \; in \; v_m \; invest \; in \; v_n \end{cases} \quad (2)$$

Furthermore, structural positionalities assess the importance of startups in the network by considering their connections. We measure structural positionalities using betweenness centrality, which captures intermediation capacity, and closeness centrality, which reflects the average distance to other nodes. Specifically, betweenness centrality quantifies the frequency with which a startup appears on the shortest paths between indirectly connected startups, indicating its ability to control information flow within the network. This can be calculated by Equation 3:

$$B(u) = \sum_{j \neq i \in V} \frac{\delta_{ij}(u)}{\delta_{ij}} \quad (3)$$

where V represents the set of nodes, δ_{ij} denotes the total number of shortest paths between nodes i and j, and $\delta_{ij}(u)$ signifies the shortest path from node i to node j that passes through node u.

Closeness centrality quantifies a startup's proximity to all other startups in the network by measuring the average length of the shortest paths from the startup to all other startups. It reflects the relative position of a node within the network, indicating how close the startup is to other startups on average. The closeness centrality measure $C(u)$ of startup u is written as Equation 4:

$$C(u) = \frac{n-1}{\sum_{v \in V \setminus \{u\}} d(u,v)} \quad (4)$$

where V represents the set of nodes, $n = |V|$ is the total number of nodes, and $d(u,v)$ denotes the shortest path distance between nodes u and v.

4.2 Temporal investment information network analysis

In the static analysis, the dynamic characteristics of the network are ignored, and only the final structure of the network is considered. However, the frequent occurrence of investment activities leads to continuously changing potential connections between startups, and these changes have a significant impact on the startups. Therefore, we construct a temporal investment information network for energy startups, maintaining the same node and edge definitions as in the static network but incorporating time information labels. In this section, we analyze the dynamic process of the network to capture the changes in investment information of startups across different periods. Network topology features are measured based on the volume of connections and structural positions [19]. We consider the relationship between a startup's access to investment and its connections with other startups as well as its position within the network structure, through temporal betweenness centrality and temporal closeness centrality.

Figure 2 visualizes the global energy startups investment information network, where each node represents a startup, and the connection between nodes represents that two startups have a common investor. When an investor invests in different startups successively, then the network has a directed edge from the former startup to the latter startup.

Mathematically, following [75],the temporal network $G(V,E,T)$ is defined at time $T = \{1,2,...,t\}$, where the edges are represented by $E = \{e^1, e^2, ..., e^t\}$.At time i, if an investor first invests in startup m and then in startup n, a directed edge e^i_{mn} is created from m to n. The weight of the edge represents the number of investors who meet this condition. For example, if investors a, b, c, and d invest in startups A and B at time i, with A being invested in before B, then the directed edge $e^i_{AB} = 4$, specifically expressed as Equations 5, 6:

$$e^t = \sum_{i=1}^{t} \sum_{m \neq n \in V} e^i_{mn} \quad (5)$$

$$e^i_{mn} = \begin{cases} 0 \; if \; there \; is \; no \; investor \; between \; v_m \; and \; v_n \; at \; time \; i \\ w^i_{mn} \; if \; investor \; in \; v_m \; invest \; in \; v_n \; at \; time \; i \end{cases} \quad (6)$$

Both temporal betweenness centrality and temporal closeness centrality require the definition of shortest paths over time. We first identify the shortest paths at each time point and then count the number of shortest paths within a specified time period. The process is described as follows in Equation 7:

$$\delta^t_{ij} = \sum_{i=1}^{t} \min_i dist_{ij} \quad (7)$$

Where $\min_i dist_{ij}$ represents the shortest path between nodes i and j at time t, and δ^t_{ij} represents the count of the shortest paths from node i to node j at each time point up to time t. Thus, temporal betweenness centrality and temporal closeness centrality can be expressed as Equations 8, 9:

$$tB(u) = \sum_{i \neq j \in V} \frac{\delta^t_{ij}(u)}{\delta^t_{ij}} \quad (8)$$

$$tC(u) = \frac{1}{\sum_{v \neq u \in V} \delta^t_{uv}} \quad (9)$$

Where $\delta^t_{ij}(u)$ represents the number of shortest paths passing through node u between nodes i and j at time t, and δ^t_{uv} represents the number of shortest paths between nodes u and v at time t.

4.3 Construction of the success evaluation model

To explore the relationship between the network variables formed by the flow of investment information among startups and the success of energy startups, we introduce Pearson correlation into our study (43) to measure the correlation of the network variables. The matrix of the Pearson correlation coefficient, P, is defined as follows in Equations 10, 11:

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mn} \end{bmatrix} \quad (10)$$

$$p_{mn} = \frac{\sum_{i=1}^{t} \left(x^m_i - \overline{x^m}\right)\left(y^n_i - \overline{y^n}\right)}{\sqrt{\sum_{i=1}^{t} \left(x^m_i - \overline{x^m}\right)^2} \sqrt{\sum_{i=1}^{t} \left(y^n_i - \overline{y^n}\right)^2}} \quad (11)$$

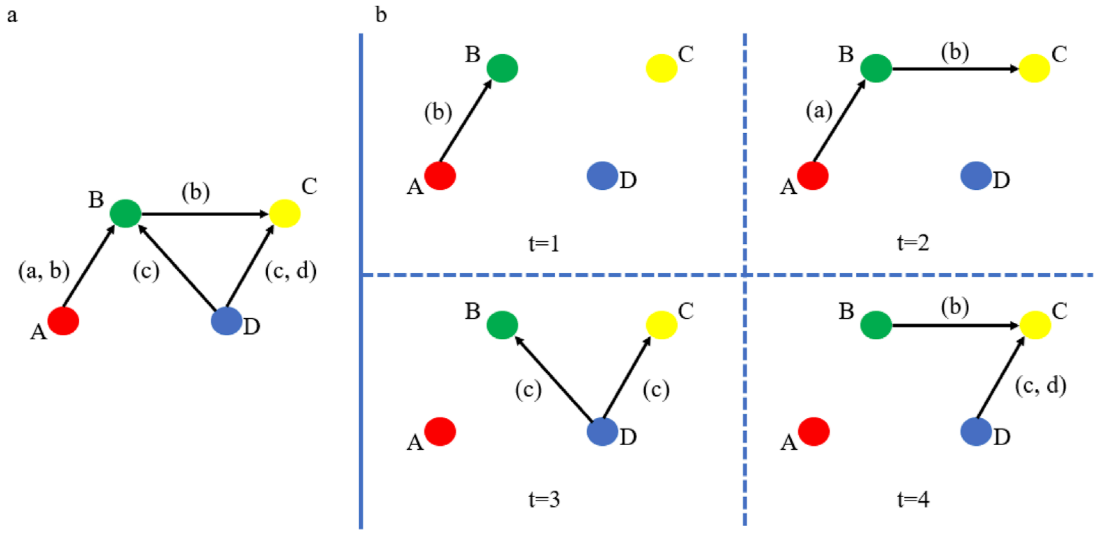


FIGURE 2
Schematic representation of global energy startups investment information network. **(a)** Static investment information network. **(b)** Temporal investment information network. Note: ABCD represent different startups, abcd denote different investors, and 1234 represent the time. The connecting edges between two startups suggest that they have a common investor, and the arrow direction represents the sequential order of the investor's investments, with the earlier-invested startup pointing to the later-invested one. The temporal network can depict the dynamics of market changes, such as the static network failing to display the investment changes of investor b due to the absence of its duration, and such changes have an impact on startups ABC.

where m and n are the length of two data series, p_{mn} is the pearson correlation coefficient between m and n, t is the number of spot data for each variable. $\overline{x^m}$ is the average percentage value of m, $\overline{y^n}$ is the average percentage value of n. We then estimate the following Equation 12 based on the relationship between the network variables and startup success:

$$Y_i = \alpha D_i + \beta B_i + \gamma X_i + \varepsilon \tag{12}$$

Where Y_i measures the success potential of startup i. α, β and γ are the weight of different scores, and ε is an adjustment factor. D_i represents the linkage volumes of startup i in the network, consisting of in-degree, out-degree, weighted in-degree, and weighted out-degree. B_i denotes the structural position of startup i in the network, including temporal betweenness centrality and temporal closeness centrality. Additionally, X_i represents control variables that may influence startup success, including government effectiveness and regulatory quality.

Finally, we evaluate the performance of the prediction model using the Area Under the Receiver Operating Characteristic Curve (AUC). Specifically, we adopt the roc_auc_score function from the Scikit-learn library [76], which is based on the Wilcoxon-Mann-Whitney statistic. This method measures the model's ability to distinguish between positive and negative samples by comparing the ranking of predicted scores. We randomly select 80% of the innovative energy startups as the training set and the remaining 20% as the testing set. The model's predictive accuracy is assessed by evaluating its ability to correctly predict the success of startups in the testing set, as shown in Equation 13:

$$AUC = \frac{1}{mn} \sum_{i \in Positive} \sum_{j \in Negative} I\left(p_i > p_j\right) \tag{13}$$

Where m is the number of positive samples, n is the number of negative samples, p_i and p_j are the predicted scores for positive and negative samples respectively, and $I\left(p_i > p_j\right) = 1$, 0.5 if $p_i = p_j$, and 0 otherwise.

5 Results and discussion

5.1 Statistic topology on investment information network

The investment information network of energy startups exhibits complex statistical topological characteristics, with varying positions and influence of different types of startups within the network. Based on static topological analysis methods, we conduct an empirical study of the investment information networks of innovative energy, renewable energy, and traditional energy startups.

Table 2 provides the basic characteristics of energy market startups in the static network. Market share indicates the percentage of the total investment funds accounted for by the startups. Closeness centrality and betweenness centrality are calculated based on the results of the centrality experiments [77].

Innovative energy startups account for less than 1% of the market, while traditional energy startups, with a market share of over 60%, remain the main force in the energy investment and financing market. However, for the three types of energy startups—innovative energy, renewable energy, and traditional energy—both the in-degree and out-degree have gradually decreased, while the weighted degree has gradually increased. Although innovative energy startups have fewer connections with

TABLE 2 Static network characteristics.

Type	Market occupancy	Average in-degree	Average out-degree	Average weighted in-degree	Average weighted out-degree	Static closeness centrality	Static betweenness centrality
Innovative Energy	0.9351%	**1.715**	**1.701**	5.365	5.350	0.118	**3.040**
Renewable Energy	30.855%	1.662	1.664	6.718	6.716	0.130	2.942
Traditional Energy	68.209%	1.564	1.564	**7.576**	**7.577**	**0.133**	2.564

Note: Bold values indicate the best performance in each column.

other startups, the weight of these connections in the network is larger. This indicates that compared to traditional and renewable energy startups, innovative energy startups may be more inclined to select specific and strategically significant partners for in-depth cooperation, such as preferring to collaborate with institutions like venture capital and private equity, which typically provide larger amounts of investment. The betweenness centrality further reflects this, as innovative energy startups in the network are more reflective of the flow trend of investment information, and startups influence the direction and speed of the dissemination of investment information. However, the closeness centrality of innovative energy startups is the lowest among the three types of startups, possibly due to their low market share, which prevents startups from reaching more startups quickly and slows access to market information.

Although innovative energy startups have a small market share, as indicated by centrality, their influence is not insignificant. Startups with high centrality are highly connected nodes in the investment and financing network, and the resulting market influence cannot be solely determined by market share. These innovative energy startups with high centrality have a huge market influence that can have a wide impact on the investment trends of the whole investment and financing and thus dominate the development of the energy industry. In addition, markets are sensitive to volatile information, and technological innovations in the energy industry and uncertainties in economic policies can exacerbate the volatility of investment and financing in the energy market, which may in turn affect investment behavior. Therefore, analyzing the investment and financing patterns of the energy market solely based on market share and the static network is restrictive. It is necessary to conduct a dynamic analysis of the investment and financing network.

5.2 Dynamic roles of innovative energy startups

We observe the development of the energy market in recent years from a dynamic perspective. Innovative energy startups entered the market in 2010. Figure 3 shows the temporal betweenness and closeness centrality values of innovative energy startups in the temporal network from 2010 to 2024. We can observe that the distribution of betweenness centrality values is relatively uniform, with the majority being 0.000356 and 0.000178. While the closeness centrality values are relatively dispersed but concentrated within the range of 0.0006–0.001. Figure 4 separately presents the temporal betweenness and closeness centrality of the three types of startups - innovative energy, renewable energy, and traditional energy - from 2005 to 2024. In Figure 4a, the peak of connectivity in 2010 represents that innovative energy startups received significant attention immediately upon entering the market. Due to political unrest in the Middle East and the shale gas revolution in the United States, the energy market declined continuously until the Paris Climate Agreement in 2015 revitalized the energy market and maintained a relatively stable development in the subsequent years, especially with innovative energy startups becoming a key center for investment. However, both the COVID-19 pandemic in 2019 and the Russia-Ukraine war in 2022 affected the global financial market, and the energy investment and financing market was also impacted due to the interrelationships of investments among startups.

We found that temporal closeness centrality differs from the traditional closeness centrality observed in static networks. Notably, with the exception of the initial phase when innovative energy startups enter the market, the three categories of startups exhibit near-total synchronization in subsequent years. As illustrated in Figure 4b, innovative energy startups displayed elevated closeness centrality during their market entry phase, a phenomenon that can be attributed to robust governmental and startup support, policy incentives, and significant capital investment. These factors collectively facilitated the rapid advancement of innovative energy technologies, bolstered investor confidence, and reinforced the startups' prominence within the network. While traditional energy sources continue to dominate global energy supply and renewable energy sectors experience rapid growth, innovative energy startups have secured a distinct strategic position within the energy network, driven by their potential for technological disruption and their capacity to shape future industry developments.

Moreover, our analysis reveals a high degree of consistency in the changes of betweenness centrality between innovative energy startups and renewable energy startups, although innovative energy startups exhibit greater volatility. This consistency can be attributed to the shared focus of both types of startups on sustainability,

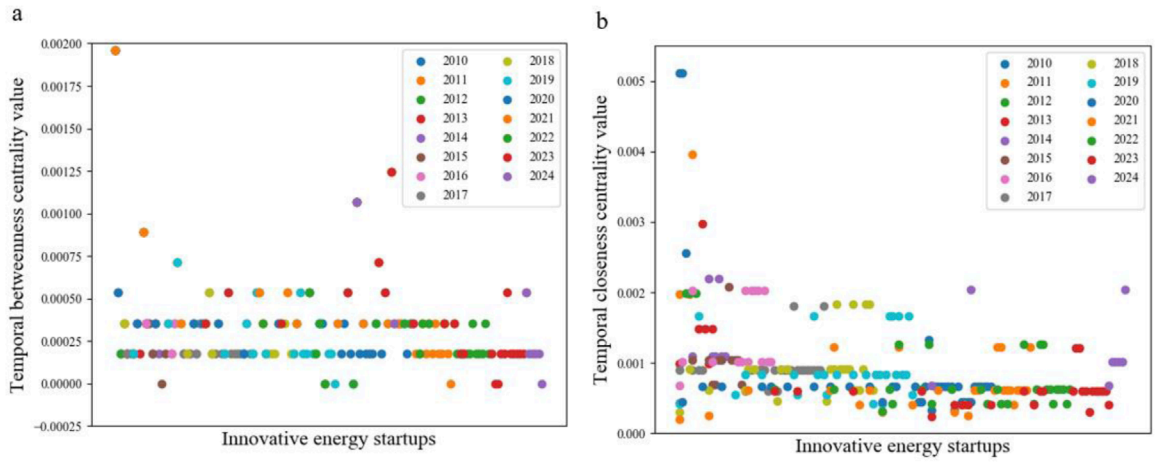


FIGURE 3
Temporal centrality values of innovative energy startups. **(a)** Temporal betweenness centrality of innovative energy startups. **(b)** Temporal closeness centrality of innovative energy startups.

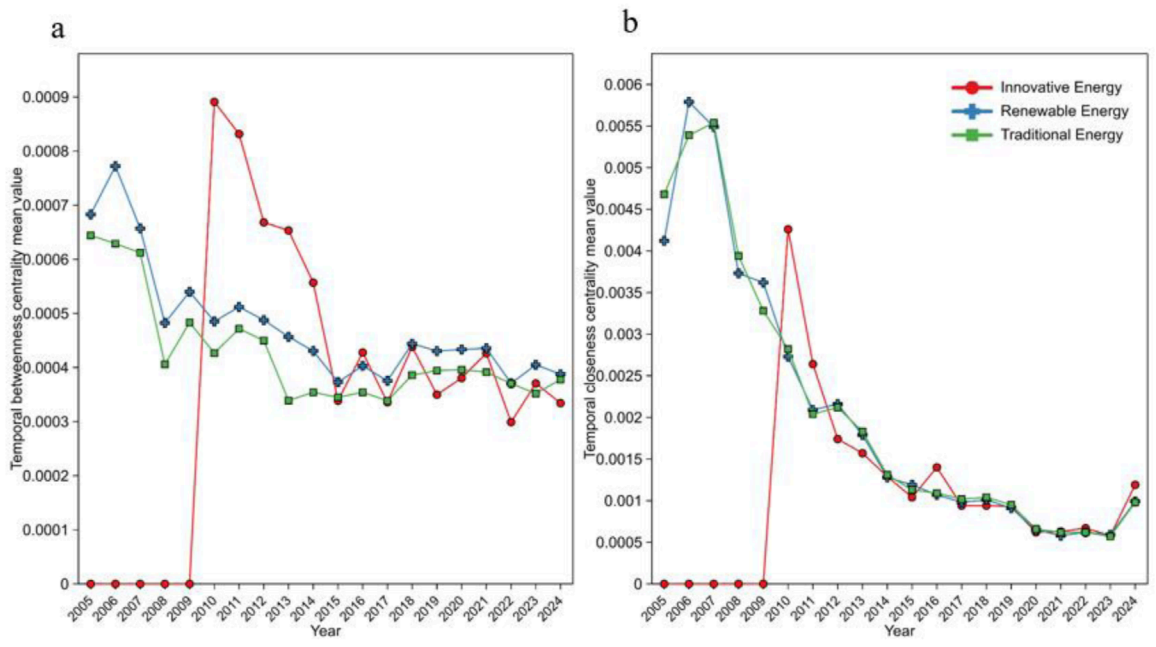


FIGURE 4
Temporal centrality of energy startups divided by industries. **(a)** Temporal betweenness centrality of startups. **(b)** Temporal closeness centrality of startups. Note: We depict the changing trends of centrality for startups classified by business type. The vertical axis of **(a)** represents the average temporal betweenness centrality, and the vertical axis of **(b)** represents the average temporal closeness centrality. Innovative energy startups are represented by red circles, renewable energy startups by blue crosses, and traditional energy startups by green squares.

their commitment to developing environmentally friendly solutions, reducing reliance on traditional energy sources such as fossil fuels, and fostering the green economy. Additionally, both categories benefit from supportive policies enacted by governments and international organizations aimed at advancing the application of and investment in clean energy technologies. However, while both innovative and renewable energy startups must remain agile in response to market dynamics, adjusting business models and technological innovations to satisfy consumer demands, innovative energy startups face heightened demands in terms of technological advancements and real-time market responsiveness. Consequently, they are more sensitive to market fluctuations. This increased sensitivity is further compounded by the fact that innovative energy startups span innovations across both traditional and renewable energy sectors, meaning that shifts in either domain can affect their market positioning. Given that innovative energy startups

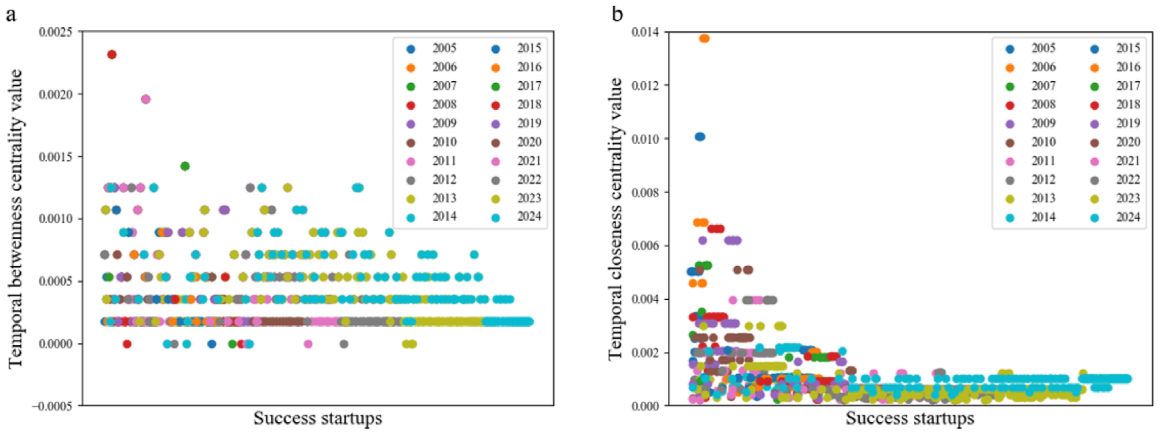


FIGURE 5
Temporal centrality values of successful startups. **(a)** Temporal betweenness centrality of successful startups. **(b)** Temporal closeness centrality of successful startups.

are still in the development phase, with technologies and markets continuing to evolve, the risks associated with investment returns are more pronounced. These factors collectively contribute to the heightened volatility observed in the betweenness centrality of innovative energy startups.

5.3 Dynamic roles of success energy startups

Through both static and dynamic analyses of the energy market, we have elucidated the characteristics of the overall energy market landscape. To investigate the relationship between successful energy startups and the energy market, we employed a temporal network framework to capture the structural positions of successful energy startups within the network, as illustrated in Figures 5, 6. Figure 5 presents the temporal trends of betweenness and closeness centrality for successful energy startups from 2005 to 2024. Our findings reveal that the distribution patterns of successful startups exhibit a notable consistency with those of innovative energy startups with regard to the distribution of betweenness centrality, which demonstrates a more uniform spread compared to closeness centrality. In the subsequent analysis, we compared the centrality values of successful startups against those of unsuccessful startups, with the results presented in Figure 6.

Through Figure 6, we find that successful startups generally have higher betweenness centrality than unsuccessful startups, while closeness centrality shows the opposite. Successful startups, as "bridge" nodes, play an important intermediary role in the flow of investment information between different startups, acting as links connecting multiple startups through shared investors. The high betweenness centrality of successful startups indicates their ability to attract a large number of investors and foster strong connections between investors, demonstrating high 'investment accessibility'. High betweenness centrality also suggests that successful startups occupy a strategic position in the energy network, holding influence and power within the entire investment ecosystem. When a startup is involved in multiple investment paths, it can control the speed of information flow or selectively distribute information and opportunities to other startups. Thus, successful startups can obtain important resources through their investment information networks and may even influence the choice of cooperation partners by investors. Low closeness centrality indicates that successful startups are more distantly connected to other startups in the network, located at relatively peripheral positions, and cannot quickly access investment resources. Although successful startups play an important bridging role in the investment network, they are not located in the core areas. This could mean that successful startups are more involved in decentralized investment fields or maintain connections with specific investment circles that are not at the center of the entire investment network. Their low closeness centrality might reflect their focus on specific fields or markets rather than broad participation in all investors' views.

Successful startups have strong cross-domain resource integration capabilities. Due to fewer direct connections with other startups, successful startups may rely on other startups to transmit information and resources, potentially affecting the timeliness of information acquisition. However, successful startups have a diverse group of investors who invest in various fields and startups, enabling them to establish extensive indirect connections through investors and access resources and information from different areas. Therefore, successful startups can acquire diverse resources and opportunities, promoting cross-domain collaboration. We find that for energy startups, although the lack of direct connections with other startups reduces the timeliness of information acquisition, increasing competition in specific fields, the diversified investment structure and indirect connections between startups may offer more flexibility and opportunities. Startups should focus on developing diversified investment portfolios to attract a broad range of investors and increase indirect connections with other startups, rather than relying solely on core investors or mainstream investment circles in the market.

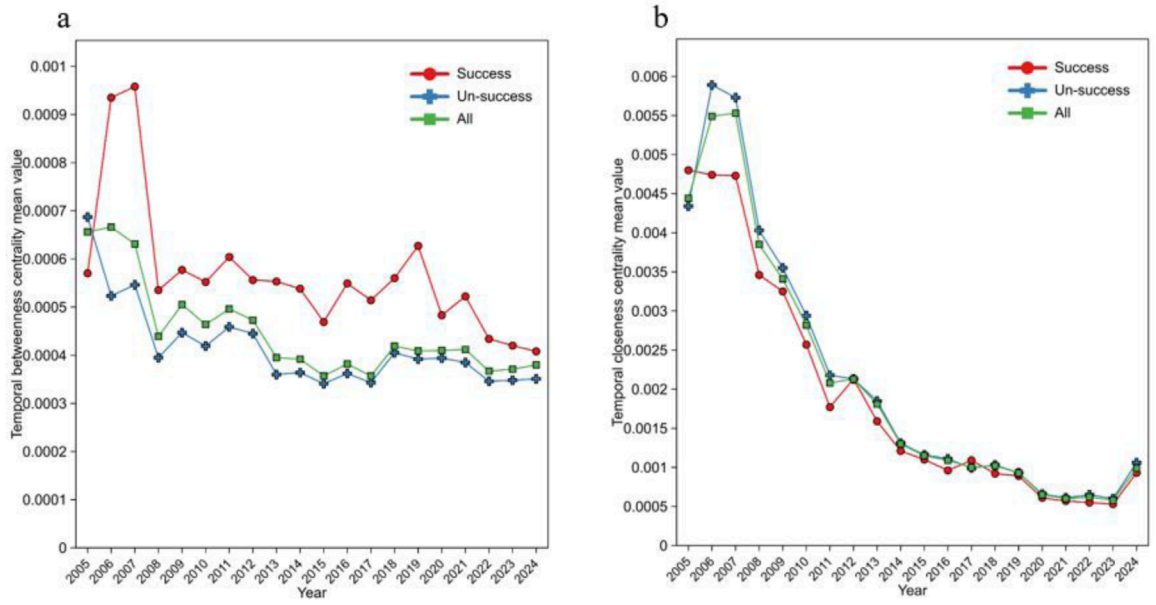


FIGURE 6
Temporal centrality of energy startups divided by success. **(a)** Temporal betweenness centrality. **(b)** Temporal closeness centrality. Note: The evolving trends in centrality for startups, categorized by success and unsuccess, are presented from 2005 to 2024. The vertical axis of **(a)** corresponds to the average temporal betweenness centrality of the startups, while the vertical axis of **(b)** reflects the average temporal closeness centrality. Successful startups are denoted by red circles, unsuccessful startups by blue crosses, and all startups are represented by green squares.

5.4 Evaluation on success potential of innovative energy startups

5.4.1 Connection between startups role characteristics and success

Table 3 presents the estimated results of the relationship between startups' combined characteristics, including their own attributes and network topological features, and startups success. Given that different network indicators conceptually capture different aspects of network role attributes, we introduce them separately into the econometric model. This mitigates issues of over-identification and multicollinearity [78]. Therefore, we estimate six independent econometric regression models, each focusing on a single network variable.

For network variables, both in-degree and weighted out-degree are statistically significantly positively correlated with successful startups at the 10% level. Additionally, out-degree, temporal betweenness centrality, and temporal closeness centrality show significance at the 5% level. These results suggest that network structure characteristics can be used to describe the characteristics of investment information in successful energy startups.

Model 1 and model 4 demonstrate relatively higher adjusted R-squared values (0.411 and 0.417, respectively), suggesting that in-degree and weighted in-degree can explain variations in startup success. This implies that the sheer number and the strength of incoming investment links may reflect external recognition or trust, which is important for startup credibility.

Regulatory quality exhibits a consistent and statistically significant positive association with startup success in five of six models, highlighting the importance of institutional environment. Government effectiveness also shows marginal significance in models, further emphasizing the role of governance in enabling entrepreneurial success.

Although model 5 and model 6 have relatively low adjusted R-squared values, the individual coefficients remain informative, suggesting that betweenness centrality and closeness centrality, while not sufficient alone, may still exert meaningful influence when considered jointly with others.

5.4.2 Accuracy of the success evaluation model

Based on the connections in 5.4.1, we design the success evaluation model, assign weights to the indicators, and evaluate its performance. Since a higher AUC indicates greater evaluation accuracy for successful startups, Table 4 confirms the validity of the proposed model. We adjusted the weights of different subsections, and notably, considering any subsection alone did not provide the best representation of the success evaluation score for successful startups. To maximize the AUC value, we exhaustively search all combinations of the weights α, β, γ, under the constraint that $\alpha+\beta+\gamma=1$, where $\alpha,\beta,\gamma \in [0,1]$ with a step size of 0.1. A higher weight indicates a greater contribution of the corresponding factor to the AUC, and thus a stronger influence on the evaluation of startup success. A weight of 0 implies that the factor has no impact on the success evaluation, while a weight of 1 indicates that the evaluation is solely determined by that single factor. When considering the network in combination with government control capacity as additional information, the AUC reaches its

TABLE 3 Regression results of startups role characteristics.

Characteristic	1	2	3	4	5	6
Indegree (log)	0.411*** (0.531)					
Outdegree (log)		0.349** (0.547)				
Weight Indegree (log)			0.417* (0.733)			
Weight Outdegree (log)				0.633*** (0.749)		
Betweenness centrality (log)					0.484** (0.627)	
Closeness centrality (log)						0.249** (0.168)
Country (log)	0.111 (0.542)	0.204 (0.385)	0.221 (1.568)	0.371* (0.644)	0.437*** (0.436)	0.382** (0.182)
Government effectiveness (log)	0.312* (1.382)	0.005 (0.843)	0.381* (1.568)	0.213 (0.919)	0.061 (2.955)	0.165* (0.745)
Regulatory quality (log)	0.752*** (1.372)	0.377** (0.843)	0.822** (1.568)	0.631*** (0.919)	0.667*** (2.954)	0.166 (0.745)
Adjusted R-squared	0.411	0.035	0.417	0.063	0.0085	0.025

Note: Standard errors are in parentheses. ***$p < 0.01$,**$p < 0.05$,*$p < 0.1$. We log-transform all variables to produce normally distributed model residuals. Additionally, before the log transformation, we add the mean to variables with zero values to address the issue of zeros in the dataset.

TABLE 4 Model precision measured by AUC.

α	β	γ	Success AUC
1	0	0	0.916
0	1	0	0.967
0	0	1	0.889
0.3	0.6	0.1	**0.984**

Note: α, β, and γ are the coefficients measuring enterprise degree, centrality, and government administrative capacity. Bold values indicate the best performance.

TABLE 5 Five-fold cross-validation results of AUC.

α	β	γ	mean_AUC	Std
0.3	0.6	0.1	0.9839	0.0052
0.4	0.6	0.0	0.9837	0.0051
0.4	0.5	0.1	0.9829	0.0056
0.2	0.7	0.1	0.9819	0.0053
0.2	0.6	0.2	0.9815	0.0060

maximum, and the network structure characteristics contribute the most to the success evaluation of startups (with high β values when maximum AUC is achieved). Therefore, considering each startup's role and dynamic trends in the network helps achieve a higher and more stable AUC (0.984), indicating better potential to assess future success.

To mitigate the risk of overfitting caused by the relatively small sample size of innovative energy startups, this study incorporated a larger dataset. Specifically, we used data from 8,900 energy startups, including 1,507 successful and 7,393 unsuccessful cases. A stratified random sampling approach was employed, allocating 80% of the data for training and 20%. When the weight parameters satisfied the condition $\alpha = 0.3, \beta = 0.6, \gamma = 0.1$, the model achieved an AUC of 0.875 on the test set, indicating a certain level of robustness.

To validate the robustness of our experiments, we employ a five-fold cross-validation approach. For each parameter combination (α, β, γ), the data is divided into five parts: in each iteration, one part is used for validation and the remaining four for training. We then compute the average AUC and its standard deviation for each combination. The top five AUC results obtained from the experiments are shown in Table 5. The standard deviation for each group is around 0.005, indicating the experiment demonstrates reasonable stability and robustness.

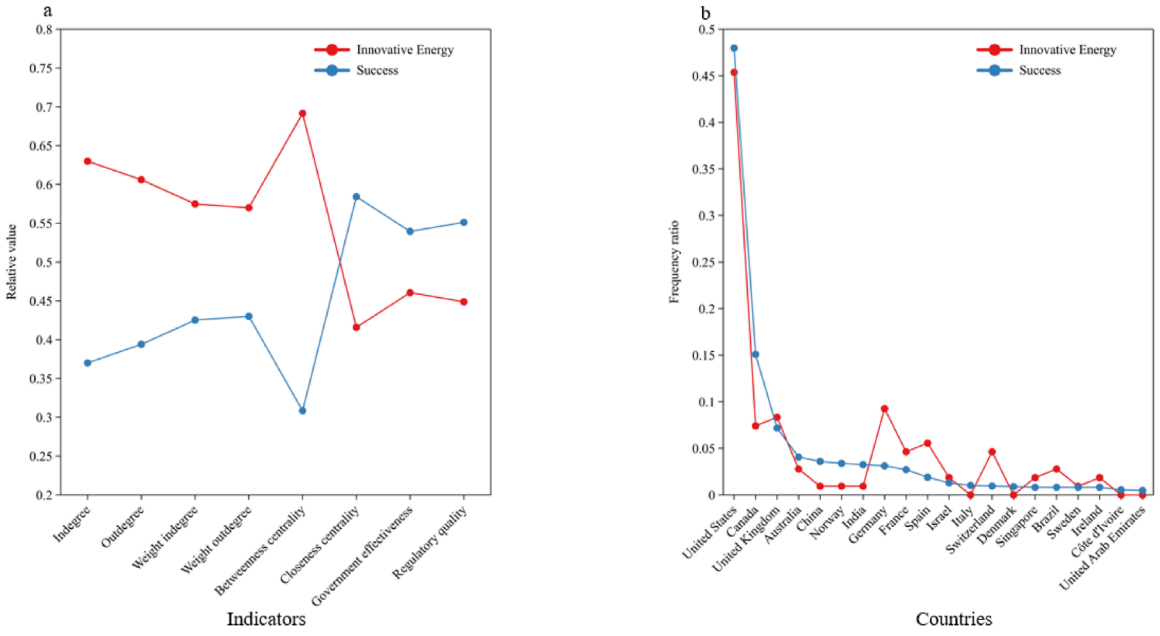


FIGURE 7
Comparison of successful energy startups and innovative energy startups. Note: We compared the values of various indicators for successful and innovative startups, as well as their national distribution. **(a)** shows the relative value (sum equal to 1) of each indicator for the startups. **(b)** shows the frequency distribution of startups in the top 20 countries where successful startups appear most frequently. Red represents innovative energy startups, and blue represents successful startups.

5.5 Estimation on success potentials of innovative energy startups

We used the success evaluation model to evaluate innovative energy startups, comparing the top 20 innovative energy startups with successful startups, as shown in Figure 7.

We found that innovative startups have higher in-degree, weighted in-degree, and betweenness centrality than successful startups, with the largest difference being in betweenness centrality. However, innovative startups lag behind successful startups in terms of closeness centrality, government efficiency, and regulatory quality.

The higher in-degree and weighted in-degree of innovative startups indicate they can attract more investment but are less effective in preventing investment outflow. The narrowing gap in weighted in-degree suggests that the strength of connections between startups plays a balancing role. Among the structural positions, betweenness centrality shows the greatest deviation between innovative and successful startups. Innovative startups have high betweenness centrality because they act as intermediaries for investment information, attracting a lot of investment. However, due to the high demands for technological updates and the high volatility of technology, compared to other startups, innovative startups' risks are greater and cannot stabilize the investment. The lower closeness centrality means innovative startups are relatively distant from others in the network, while they play an important role in connecting different startups, they are not located in the network's central area and cannot quickly reach other startups, possibly being far from the core decision-making circle of investments. Energy startups can play a controlling role in the flow of information, but their efficiency in actively acquiring and disseminating information is relatively low. They rely more on other startups to obtain investment information, and sending investment information to other parts of the network takes more time. Innovative energy startups play key roles in certain paths or network branches, ensuring the connectivity of different sub-networks or startup groups, but their influence or efficiency is more localized in specific sub-networks and not in the information center of the entire network.

Innovative startups perform worse than successful startups in terms of government efficiency and regulatory quality, indicating that government administrative capacity has a positive impact on promoting startup development. This emphasizes the critical role of high-quality, efficient services and a supportive regulatory environment in driving startup progress [19]. Innovative startups and successful startups show similarities in the countries they are located in, especially in developed countries like the U.S. and the U.K. On the one hand, emerging startups have the advantage of catching up by rapidly adopting innovative technologies in their industrial structure, and developed countries typically have more advanced technologies that startups can learn from to gain advantages [79]. On the other hand, governments in developed economies have been actively implementing policies to accelerate emerging technologies like AI and blockchain, and startups use effective policies to speed up technological innovation and follow or even lead market trends. For example, the Inflation Reduction Act enacted in

the United States in 2022 (https://www.govinfo.gov/content/pkg/PLAW-117publ169/html/PLAW-117publ169.htm) provides targeted tax incentives for energy-related enterprises, including renewable energy development, electric vehicle adoption, and energy efficiency improvements. This policy significantly lowers the financing barriers for relevant startups and enhances their survivability and competitiveness during the early stages of development. Therefore, the country and its administrative capacity have a positive impact on energy startups, playing a key role in promoting their investment.

6 Conclusion

To achieve the energy development goals, the future development of innovative energy startups has attracted active exploration from global businesses and governments. In this paper, we contribute to this discussion by analyzing the positioning of innovative energy startups in the market through the construction of an investment information network among energy startups and by evaluating the future development directions of these startups based on assessments of their success. We argue that, in addition to the influence of domestic government administrative capabilities and policies, the development of innovative energy startups is also shaped by various forms of role attributes within the investment information network.

Firstly, although innovative energy startups account for only 0.9351% of the global energy market, their high centrality in the network suggests that they serve as highly connected nodes within the investment information network. Thus, market influence does not solely depend on market share; these highly central innovative energy startups, possessing significant market influence, can exert widespread effects on the investment trends in financing, thus shaping the development of the energy industry. From the perspectives of in-degree and weighted in-degree, innovative energy startups are more likely to engage in deep collaborations with specific, strategically important partners, and these deep collaborators typically provide significant amounts of investment.

Secondly, when innovative energy startups first enter the market, they attract substantial attention, aided by policy incentives and significant capital inflows. The connectivity of these startups fluctuates frequently, which can be explained by recent social and extreme economic events. The volatility of innovative energy startups is consistent with that of renewable energy startups during the same period, likely due to their shared focus on sustainability and benefit from green economy policies. However, innovative energy startups are more pronounced than renewable energy startups, being more sensitive to technology and market potential, and tend to react more strongly to market fluctuations.

Thirdly, successful energy startups exhibit high betweenness centrality and low closeness centrality within the network compared to other startups. High betweenness centrality indicates the ability to control the speed of information flow or selectively allocate information or opportunities to other startups, while low closeness centrality suggests that these startups are more inclined to participate in decentralized investment areas or non-network-centered investment circles. Thus, successful startups possess strong cross-domain resource integration capabilities, with a diversified investor base leading to extensive indirect connections between startups. As a result, these startups can access a wide range of resources and opportunities despite having fewer direct links with other startups.

Subsequent quantitative analysis of various indicators of successful startups reveals that network degree, network centrality, and government administrative capacity all have an impact on startup success. To further quantify the contribution of each indicator, we conducted an in-depth analysis using the success evaluation model, which has been proven to be reliable. The results indicate that network structural characteristics contribute the most to the evaluation of successful startups, with the model achieving an AUC of 0.967 when the weight is set to 1. Further optimization of weight allocation shows that when network degree, network centrality, and government administrative capacity are assigned weights of 0.3, 0.6, and 0.1, respectively, the AUC of the model improves to 0.984, demonstrating that this weighting scheme more effectively assesses startup success potential.

Finally, through the analysis of successful startups, we find that innovative energy startups can enhance their competitiveness and attractiveness by improving their market layout, business models, and strategies, gaining more capital favor and improving their potential for success and development opportunities. First, a high in-degree indicates that innovative startups have the ability to attract investment, but it is more important to stabilize existing investments. Secondly, the high betweenness centrality and low closeness centrality of innovative energy startups indicate they have the ability to control the flow of investment information, but their information dissemination efficiency is low. Therefore, startups need to expand their influence beyond localized or specific sub-networks and enhance their ability to actively acquire and disseminate investment information, thereby positioning themselves in the network's center. Central positioning in capital flows provides startups with opportunities to access diverse resources, capabilities, and markets [80], creating excellent opportunities for knowledge sharing and learning [81], which have strategic value, allowing startups to acquire new technologies before widespread adoption, creating a virtuous cycle in technological innovation. Finally, in addition to developed countries, startups should also consider countries with proactive policies and clear development directions. For example, the Chinese government released the "Energy Production and Consumption Revolution Strategy (2016-2030)" in 2016 (https://www.gov.cn/xinwen/2017-04/25/content_5230568.htm), clearly stating the need to promote the deep integration of the internet with distributed energy technology, advanced grid technology, and energy storage technology, and to vigorously develop smart energy technologies. In 2020, the "Energy Development in the New Era" (http://www.scio.gov.cn/gxzt/dtzt/2020/xsddzgnyfzbps/) further transformed the government's functions, supporting the development of new technologies, models, and industries, and promoting the deep integration of energy technology with modern information, materials, and advanced manufacturing technologies. China is advancing the energy industry toward intelligence and greening

through policy support, market opening, innovative applications, digital transformation, and achieving carbon emission goals.

Therefore, based on the results of this study, governments should support high-potential innovative energy startups, particularly in their early stages, through instruments such as guidance funds and investment subsidies. These measures can help stabilize existing investor relationships and reduce financing uncertainty. In addition, for startups that serve as bridging nodes in the network, governments may consider designating them as key channels for technology diffusion by offering targeted support through technology pilot programs and public service platforms, thereby enhancing the overall efficiency of industry innovation. From the startup perspective, startups should actively improve their capabilities in acquiring and disseminating information. This can be achieved by participating in industry alliances and advancing digital infrastructure to strengthen their learning and communication capacities. Furthermore, they should seek to connect with diverse investor groups to expand their influence within the network. In the context of globalization, startups are also encouraged to explore emerging markets with strong policy orientation and substantial support, such as China, in order to leverage policy incentives and reinforce their central positioning in global capital and innovation networks.

Our study has some limitations. First, we focus on the investment information flow between startups, without considering intra-startup relationships. Future research could integrate the internal dynamics of startups and their impact on development. Second, we use IPO and M&A status as indicators of startup success. While this approach offers clear and observable outcomes, it does not capture other important dimensions of success, such as revenue growth, market influence, or long-term sustainability. We plan to explore alternative success metrics in future work as more comprehensive data become available. Furthermore, the proportion of innovative energy startups in our sample is relatively low. This distribution reflects the actual scarcity of such startups within the broader energy startup ecosystem, but it may also limit the generalizability of our findings regarding innovative energy startups. Future research should incorporate larger and more representative samples of innovative energy startups to further validate the conclusions drawn in this study. Finally, this study finds a significant association between high centrality and startup success, but does not establish a causal relationship. Future research may adopt longitudinal approaches such as panel data models, which track the characteristics of multiple entities over time, to examine the causal mechanisms implied by our findings.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YW: Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft. QG: Conceptualization, Funding acquisition, Supervision, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors are grateful for the funding from the National Natural Science Foundation of China (Grant No. 42001236), Young Elite Scientist Sponsorship Program by Bast (Grant No. BYESS2023413), and the National Natural Science Foundation of China (Grant No. 71991481, 71991480).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Tu Y, Wu W. How does green innovation improve enterprises' competitive advantage? The role of organizational learning. *Sustainable Prod Consumption* (2021) 26:504–16. doi:10.1016/j.spc.2020.12.031

2. Ecker F, Hahnel UJJ, Spada H. Promoting decentralized sustainable energy systems in different supply scenarios: the role of autarky aspiration. *Front Energy Res* (2017) 5:14. doi:10.3389/fenrg.2017.00014

3. Wurlod JD, Noailly J. The impact of green innovation on energy intensity: an empirical analysis for 14 industrial sectors in OECD countries. *Energy Econ* (2018) 71:47–61. doi:10.1016/j.eneco.2017.12.012

4. Rehman SU, Kraus S, Shah SA, Khanin D, Mahto RV. Analyzing the relationship between green innovation and environmental performance in large manufacturing firms. *Technol Forecast Social Change* (2021) 163:120481. doi:10.1016/j.techfore.2020.120481

5. Ali S, Ijaz MS, Yousaf I, Li Y. RETRACTED: connectedness and portfolio management between renewable energy tokens and metals: evidence from TVP-VAR approach. *Energy Econ* (2023) 127:107103. doi:10.1016/j.eneco.2023.107103

6. Gu Q, Li S, Tian S, Wang Y. Impact of climate risk on energy market risk spillover: evidence from dynamic heterogeneous network analysis. *Energy Econ* (2024) 137:107775. doi:10.1016/j.eneco.2024.107775

7. Uddin GS, Luo T, Yahya M, Jayasekera R, Rahman ML, Okhrin Y. Risk network of global energy markets. *Energy Econ* (2023) 125:106882. doi:10.1016/j.eneco.2023.106882

8. Boons F. *Creating ecological value: an evolutionary approach to business strategies and the natural environment*. Cheltenham, UK: Edward Elgar Publishing (2009).

9. Henry M, Bauwens T, Hekkert M, Kirchherr J. A typology of circular start-ups: an analysis of 128 circular business models. *J Clean Prod* (2020) 245:118528. doi:10.1016/j.jclepro.2019.118528

10. Ancona A, Cinelli M, Ferraro G, Iovanella A. Network-based principles of entrepreneurial ecosystems: a case study of a start-up network. *Small Business Econ* (2023) 61(4):1497–514. doi:10.1007/s11187-023-00738-6

11. Yu Y, Qayyum M, Li S. Trade dynamics of environmental goods within global energy economy and their impacts on green technological innovation: a complex network analysis. *Energy Econ* (2024) 140:107957. doi:10.1016/j.eneco.2024.107957

12. Okrah J, Nepp A. Factors affecting startup innovation and growth. *Journal of Advanced Management Science* (2018) 6(1):34–38. doi:10.18178/joams.6.1.34-38

13. Berger LA, Berger DR. *The talent management handbook: creating a sustainable competitive advantage by selecting, developing, and promoting the best people*. McGraw-Hill (2011).

14. Capasso M, Treibich T, Verspagen B. The medium-term effect of R&D on firm growth. *Small Business Econ* (2015) 45:39–62. doi:10.1007/s11187-015-9640-6

15. Bocken NMP. Sustainable venture capital–catalyst for sustainable start-up success? *J Clean Prod* (2015) 108:647–58. doi:10.1016/j.jclepro.2015.05.079

16. Skawińska E, Zalewski RI. Success factors of startups in the EU—A comparative study. *Sustainability* (2020) 12(19):8200. doi:10.3390/su12198200

17. Melcangi D, Turen J. Subsidizing startups under imperfect information. *J Monetary Econ* (2023) 139:93–109. doi:10.1016/j.jmoneco.2023.06.008

18. Awaya Y, Krishna V. Startups and upstarts: disadvantageous information in r&d. *J Polit Economy* (2021) 129(2):534–69. doi:10.1086/711953

19. Zhang J, Ballas D, Liu X. Global climate change mitigation technology diffusion: a network perspective. *Energy Econ* (2024) 133:107497. doi:10.1016/j.eneco.2024.107497

20. Xu JH, Yi BW, Fan Y. Economic viability and regulation effects of infrastructure investments for inter-regional electricity transmission and trade in China. *Energy Econ* (2020) 91:104890. doi:10.1016/j.eneco.2020.104890

21. Kennedy KM, Edwards MR, Doblinger C, Thomas ZH, Borrero MA, Williams ED, et al. The effects of corporate investment and public grants on climate and energy startup outcomes. *Nat Energy* (2024) 9:883–93. doi:10.1038/s41560-024-01530-w

22. Guan W, Guan Q, Duan Y, Xiang C. Finding successful startups by using information flows among investors in higher order network of investments. *IEEE Trans Comput Social Syst* (2024) 11(5):5803–14. doi:10.1109/tcss.2024.3394439

23. Chi Y, Hu N, Lu D, Yang Y. Green investment funds and corporate green innovation: from the logic of social value. *Energy Econ* (2023) 119:106532. doi:10.1016/j.eneco.2023.106532

24. Rehman IU, Shahzad F, Laique U, Hanif MA. Does environmental innovation improve investment efficiency. *Borsa Istanbul Rev* (2024) 24(1):164–75. doi:10.1016/j.bir.2023.11.007

25. Xu J, Zeng S, Qi S, Cui J. Do institutional investors facilitate corporate environmental innovation. *Energy Econ* (2023) 117:106472. doi:10.1016/j.eneco.2022.106472

26. Siciliano G, Cantoni R, Lonn P, Por N, Kry S, Morn C, et al. "leave no one behind". A power-capabilities-energy justice perspective on energy transition in remote rural communities in Cambodia. *World Development* (2025) 185:106793. doi:10.1016/j.worlddev.2024.106793

27. Bu Y, Wang E, Bai J, Shi Q. Spatial pattern and driving factors for interprovincial natural gas consumption in China: based on SNA and LMDI. *J Clean Prod* (2020) 263:121392. doi:10.1016/j.jclepro.2020.121392

28. Singh M, Jiao J, Klobasa M, Frietsch R. Making energy-transition headway: a data driven assessment of German energy startups. *Sustainable Energy Tech Assessments* (2021) 47:101322. doi:10.1016/j.seta.2021.101322

29. Cheng D, Niu Z, Zhang L. Delinquent events prediction in temporal networked-guarantee loans. *IEEE Trans Neural Networks Learn Syst* (2020) 34(4):1692–704. doi:10.1109/tnnls.2020.3027346

30. Zhao L, Wang GJ, Wang M, Bao W, Li W, Stanley HE. Stock market as temporal network. *Physica A: Stat Mech its Appl* (2018) 506:1104–12. doi:10.1016/j.physa.2018.05.039

31. Gallagher KS, Holdren JP, Sagar AD. Energy-technology innovation. *Annu Rev Environ Resour* (2006) 31(1):193–237. doi:10.1146/annurev.energy.30.050504.144321

32. Plank J, Doblinger C. The firm-level innovation impact of public R&D funding: evidence from the German renewable energy sector. *Energy policy* (2018) 113:430–8. doi:10.1016/j.enpol.2017.11.031

33. MacGillivray A, Jeffrey H, Wallace R. The importance of iteration and deployment in technology development: a study of the impact on wave and tidal stream energy research, development and innovation. *Energy Policy* (2015) 87:542–52. doi:10.1016/j.enpol.2015.10.002

34. Kim K, Kim Y. Role of policy in innovation and international trade of renewable energy technology: empirical study of solar PV and wind power technology. *Renew Sustainable Energy Rev* (2015) 44:717–27. doi:10.1016/j.rser.2015.01.033

35. Darmani A, Arvidsson N, Hidalgo A, Albors J. What drives the development of renewable energy technologies? Toward a typology for the systemic drivers. *Renew Sustainable Energy Rev* (2014) 38:834–47. doi:10.1016/j.rser.2014.07.023

36. Grubler A, Anadón LD, Gallagher KS, Kempener R, O'Rourke A, Wilson C. Global R&D, market formation, and diffusion investments in energy technology innovation. In: *Learning from historical successes and failures*. Cambridge, UK: Cambridge University Press (2013). p. 292.

37. Karltorp K, Guo S, Sandén BA. Handling financial resource mobilisation in technological innovation systems-the case of Chinese wind power. *J Clean Prod* (2017) 142:3872–82. doi:10.1016/j.jclepro.2016.10.075

38. Smith A, Contreras GAT, Brisbois MC, Lacey-Barnacle M, Sovacool BK. Inclusive innovation in just transitions: the case of smart local energy systems in the UK. *Environ Innovation Societal Transitions* (2023) 47:100719. doi:10.1016/j.eist.2023.100719

39. Reichman JH, Rai AK, Newell RG, Wiener JB. Intellectual property and alternatives: strategies for green innovation. *Intellect Property Rights Leg Econ Challenges Development* (2014) 356–91. doi:10.1093/acprof:oso/9780199660759.003.0012

40. Goldstein A, Doblinger C, Baker E, Anadón LD. Startups supported by ARPA-E were more innovative than others but an investment gap May remain. *Nat Energy* (2020) 5(10):741–2. doi:10.1038/s41560-020-00691-8

41. Jordaan SM, Romo-Rabago E, McLeary R, Reidy L, Nazari J, Herremans IM. The role of energy technology innovation in reducing greenhouse gas emissions: a case study of Canada. *Renew Sustainable Energy Rev* (2017) 78:1397–409. doi:10.1016/j.rser.2017.05.162

42. Jenkins J. *Bridging the clean energy valleys of death: helping American entrepreneurs meet the nations energy innovation imperative*. Oakland, CA: Breakthrough Institute (2022).

43. Guan Q, An H, Wang K, Duan Y, Zhang Y. Functional trade patterns and their contributions to international photovoltaic trade revealed by network motifs. *Energy* (2020) 195:116989. doi:10.1016/j.energy.2020.116989

44. Bonaventura M, Ciotti V, Panzarasa P, Liverani S, Lacasa L, Latora V. Predicting success in the worldwide start-up network. *Scientific Rep* (2020) 10(1):345. doi:10.1038/s41598-019-57209-w

45. Mann RJ, Sager TW. Patents, venture capital, and software start-ups. *Res Policy* (2007) 36(2):193–208. doi:10.1016/j.respol.2006.10.002

46. Arroyo J, Corea F, Jimenez-Diaz G, Recio-Garcia JA. Assessment of machine learning performance for decision support in venture capital investments. *Ieee Access* (2019) 7:124233–43. doi:10.1109/access.2019.2938659

47. Hegde D, Tumlinson J. Does social proximity enhance business partnerships? Theory and evidence from ethnicity's role in US venture capital. *Management Sci* (2014) 60(9):2355–80. doi:10.1287/mnsc.2013.1878

48. Nanda R, Rhodes-Kropf M. Investment cycles and startup innovation. *J Financial Econ* (2013) 110(2):403–18. doi:10.1016/j.jfineco.2013.07.001

49. Banerji D, Reimer T. Startup founders and their LinkedIn connections: are well-connected entrepreneurs more successful? *Comput Hum Behav* (2019) 90:46–52. doi:10.1016/j.chb.2018.08.033

50. Gloor PA, Colladon AF, Grippa F, Hadley BM, Woerner S. The impact of social media presence and board member composition on new venture success: evidences from vc-backed US startups. *Technol Forecast Social Change* (2020) 157:120098. doi:10.1016/j.techfore.2020.120098

51. Huang L, Pearce JL. Managing the unknowable: the effectiveness of early-stage investor gut feel in entrepreneurial investment decisions. *Administrative Sci Q* (2015) 60(4):634–70. doi:10.1177/0001839215597270

52. Maarouf A, Feuerriegel S, Pröllochs N. A fused large language model for predicting startup success. *Eur J Oper Res* (2025) 322(1):198–214. doi:10.1016/j.ejor.2024.09.011

53. Connelly BL, Certo ST, Ireland RD, Reutzel CR. Signaling theory: a review and assessment. *J Management* (2011) 37(1):39–67. doi:10.1177/0149206310388419

54. Parhankangas A, Ehrlich M. How entrepreneurs seduce business angels: an impression management approach. *J Business Venturing* (2014) 29(4):543–64. doi:10.1016/j.jbusvent.2013.08.001

55. Huggins R, Thompson P. A network-based view of regional growth. *J Econ Geogr* (2014) 14(3):511–45. doi:10.1093/jeg/lbt012

56. Huggins R, Johnston A, Thompson P. Network capital, social capital and knowledge flow: how the nature of inter-organizational networks impacts on innovation. *Industry and Innovation* (2012) 19(3):203–32. doi:10.1080/13662716.2012.669615

57. Ferrier GD, Reyes J, Zhu Z. Technology diffusion on the international trade network. *J Public Econ Theor* (2016) 18(2):291–312. doi:10.1111/jpet.12186

58. Faems D, Bos B, Noseleit F, Leten B. Multistep knowledge transfer in multinational corporation networks: when do subsidiaries benefit from unconnected sister alliances? *J Management* (2020) 46(3):414–42. doi:10.1177/0149206318798037

59. Bakshy E, Rosenn I, Marlow C, Adamic L. The role of social networks in information diffusion. In: Proceedings of the 21st international conference on World Wide Web (2012). p. 519–28.

60. Zhang ZK, Liu C, Zhan XX, Lu X, Zhang CX, Zhang YC. Dynamics of information diffusion and its applications on complex networks. *Phys Rep* (2016) 651:1–34. doi:10.1016/j.physrep.2016.07.002

61. Zhang L, Luo M, Boncella RJ. Product information diffusion in a social network. *Electron Commerce Res* (2020) 20:3–19. doi:10.1007/s10660-018-9316-9

62. Ying J. Gradual information diffusion across commonly owned firms. *J Financial Econ* (2024) 156:103852. doi:10.1016/j.jfineco.2024.103852

63. Barabási AL, Albert R. Emergence of scaling in random networks. *Science* (1999) 286(5439):509–12. doi:10.1515/9781400841356.349

64. Dong Q, Zhao Y, Ma X, Zhou Y. Risk spillover between carbon markets and stock markets from a progressive perspective: measurements, spillover networks, and driving factors. *Energy Econ* (2024) 129:107228. doi:10.1016/j.eneco.2023.107228

65. Fire M, Guestrin C. The rise and fall of network stars: analyzing 2.5 million graphs to reveal how high-degree vertices emerge over time. *Inf Process and Management* (2020) 57(2):102041. doi:10.1016/j.ipm.2019.05.002

66. Nsour F, Sayama H. Hot-get-richer network growth model. In: International Conference on Complex Networks and Their Applications. Cham: Springer International Publishing (2020). p. 532–43.

67. Barnes MR, Nicosia V, Clegg RG. Measuring social mobility in temporal networks. *Scientific Rep* (2025) 15(1):5941. doi:10.1038/s41598-025-89090-1

68. Wang Z, Li M, Lu J, Cheng X. Business innovation based on artificial intelligence and blockchain technology. *Inf Process and Management* (2022) 59(1):102759. doi:10.1016/j.ipm.2021.102759

69. Teng M, Shen M. Fintech and energy efficiency: evidence from OECD countries. *Resour Policy* (2023) 82:103550. doi:10.1016/j.resourpol.2023.103550

70. Ari I, Koc M. Philanthropic-crowdfunding-partnership: a proof-of-concept study for sustainable financing in low-carbon energy transitions. *Energy* (2021) 222:119925. doi:10.1016/j.energy.2021.119925

71. Dej D. *Defining and measuring entrepreneurial success. Entrepreneurship: a psychological approach* (2010). p. 89–102.

72. Amoroso N, Bellantuono L, Monaco A, De Nicolò F, Somma E, Bellotti R. Economic interplay forecasting business success. *Complexity* (2021) (1):8861267. doi:10.1155/2021/8861267

73. Kaufmann D, Kraay A, Mastruzzi M. The worldwide governance indicators: methodology and analytical issues. *World Bank Policy Res Working Paper* (2011) 3(2):220–46. doi:10.1017/s1876404511200046

74. Newman M. *Networks*. Oxford, UK: Oxford University Press (2018).

75. Holme P, Saramäki J. Temporal networks. *Phys Rep* (2012) 519(3):97–125. doi:10.1016/j.physrep.2012.03.001

76. Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learn* (2001) 45:171–86. doi:10.1023/a:1010920819831

77. Brandes U. A faster algorithm for betweenness centrality. *J Math Sociol* (2001) 25(2):163–77. doi:10.1080/0022250x.2001.9990249

78. Shi S, Wong SK, Zheng C. Network capital and urban development: an inter-urban capital flow network analysis. *Reg Stud* (2022) 56(3):406–19. doi:10.1080/00343404.2021.1955098

79. Grubler A. Energy transitions research: insights and cautionary tales. *Energy policy* (2012) 50:8–16. doi:10.1016/j.enpol.2012.02.070

80. Lin JY. New structural economics: a framework for rethinking development. *The World Bank Res Observer* (2011) 26(2):193–221. doi:10.1093/wbro/lkr007

81. Cheng HWJ. Factors affecting technological diffusion through social networks: a review of the empirical evidence. *The World Bank Res Observer* (2022) 37(2):137–70. doi:10.1093/wbro/lkab010

82. Díaz-Santamaría C, Bulchand-Gidumal J. Econometric estimation of the factors that influence startup success. *Sustainability* (2021) 13(4):2242.

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 07 August 2025
DOI 10.3389/fphy.2025.1647667

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Shijia Song,
Chongqing University, China
Luiz De Jesus,
Loyola Andalusia University, Spain
Difang Huang,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Li Lin,
llin@ecust.edu.cn

RECEIVED 16 June 2025
ACCEPTED 07 July 2025
PUBLISHED 07 August 2025

CITATION
Pang W-Y and Lin L (2025) Market state transitions and crash early warning in the Chinese stock market.
Front. Phys. 13:1647667.
doi: 10.3389/fphy.2025.1647667

COPYRIGHT
© 2025 Pang and Lin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Market state transitions and crash early warning in the Chinese stock market

Wu-Yue Pang[1,2] and Li Lin[1,3]*

[1]Department of Finance, Business School, East China University of Science and Technology, Shanghai, China, [2]School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China, [3]Risk-Center, ETH Zürich, Zürich, Switzerland

Understanding systemic risk in financial markets requires tools that capture both structural complexity and dynamic evolution. This study adopts a complex systems perspective to analyze the Chinese stock market through correlation-based state classification. Using multidimensional scaling and K-means clustering on rolling stock return correlations, we identify five distinct market states that reflect varying degrees of systemic co-movement and exhibit strong temporal persistence and local transition patterns. We find that transitions among these states encode meaningful information about market structural shifts and are closely linked to the emergence of crash conditions. Building on this insight, we construct an early warning model using decision trees trained on temporal features derived from market state transitions—including medium-term state distributions, directional change ratios, and short-term evolutionary paths. The model achieves high recall and precision across configurations, and supports real-time adaptability via projection-based state labeling. These results highlight the value of market state dynamics and their transitions as a basis for systemic risk monitoring and crash anticipation in complex financial systems.

KEYWORDS

market state transition, systemic risk, early warning, correlation structure, stock market

1 Introduction

Financial markets are widely regarded as complex adaptive systems, in which the interactions among heterogeneous agents give rise to emergent phenomena such as volatility clustering, contagion, and crashes [1–4]. Capturing these dynamics requires moving beyond isolated asset-level indicators to examine the structure and evolution of market-wide interactions. In this context, correlation-based and network-theoretic representations have emerged as powerful tools for modeling systemic behavior. Early studies constructed networks from return correlation matrices to quantify collective dynamics and market fragility [5, 6], while more recent work has introduced higher-order interaction structures, including Granger-causality networks [7], speculative influence networks [8], portfolio overlap networks [9], and Ricci-curvature-based geometric approaches [10, 11]. These approaches collectively highlight the importance of modeling financial markets as evolving structures whose topological and temporal features govern their stability—serving, in effect, as emergent order parameters that summarize the global configuration of the system [12, 13]. In parallel, recent advances in econometric modeling have enriched this structural perspective through dynamic correlation and network

connectivity analyses [14], contagion mechanism estimation [15], and cross-sectional uncertainty metrics [16], offering additional insights into the spatiotemporal propagation of systemic stress.

Building on these foundations, a growing body of research has demonstrated that financial markets can be decomposed into a small number of discrete market states, derived from clustering time-dependent correlation matrices [17–19]. These states are understood to reflect varying levels of market-wide coordination, with high-correlation regimes typically aligned with heightened systemic risk. Crucially, transitions between these states are not random but exhibit structured, often persistent patterns—indicative of path-dependent dynamics rooted in collective behavior and feedback effects [20, 21]. This suggests that financial instability is less a sudden shock than a gradual process, emerging from shifts in the system's internal organization and the evolution of its interaction topology. Such state transitions are also closely related to latent patterns of risk contagion and inter-market connectivity, as highlighted in recent econometric models of dynamic transmission mechanisms and time-zone-adjusted VAR frameworks [15]. Moreover, advances in modeling cross-sectional uncertainty have shown promise in enhancing correlation-based classification methods, offering insights into the predictive structure of latent regimes [16]. These perspectives collectively underscore the need to treat state transitions not merely as clustering artifacts, but as meaningful structural shifts reflecting endogenous risk propagation.

Despite these advances, several important aspects remain underexplored in the existing literature. First, most studies are retrospective in nature—focusing on identifying and interpreting market states ex post—without translating these structural insights into forward-looking monitoring tools. Second, while emerging research suggests that market state transitions encode valuable early signals of systemic stress [22, 23], relatively few studies have formally integrated such temporal patterns into operational early warning systems. This issue is especially relevant in the context of emerging markets such as China, where distinctive market microstructures from developed market and investor behaviors complicate risk assessment. Existing models often lack structural interpretability and real-time adaptability, underscoring the need for monitoring frameworks that reflect the endogenous evolution of systemic fragility [24–26].

To address the aforementioned limitations, this paper proposes a structure-aware early warning framework that explicitly incorporates the temporal dynamics of market state transitions. Using the CSI 300 Index as a representative system, we first identify five distinct market states through a combination of correlation-based similarity measures, multidimensional scaling (MDS), and K-means clustering. These states correspond to different levels of systemic coupling and exhibit empirically stable transition patterns: transitions occur predominantly between adjacent states, while high-correlation regimes are disproportionately associated with historical crash periods.

Leveraging these insights, we develop a decision tree-based early warning model constructed from temporally structured features of market state dynamics. These include medium-term state distributions, directional transition ratios, and recent evolutionary motifs. The model demonstrates high recall and precision under cross-validation and remains robust when extended to new data via an efficient projection-based labeling mechanism. Crucially, unlike traditional indicators based on instantaneous correlation structure—such as the Absorption Ratio (AR) [6] and its extension ARR [27], eigenvalue-based early warning signals [20], or curvature-derived fragility measures [10, 11]—our framework emphasizes the predictive value of how structural regimes evolve over time. This transition-centric approach provides a dynamic lens on market fragility, enhancing the model's interpretability and its potential for real-time monitoring in complex financial environments.

To the best of our understanding, few studies have systematically leveraged market state transitions—interpreted as emergent order-parameter dynamics within complex systems—to inform supervised early warning models, particularly in the context of emerging markets. This study seeks to bridge complex systems modeling with practical risk monitoring by operationalizing structurally grounded state transitions as interpretable signals for crash forecasting.

The remainder of the paper is organized as follows. Section 2.1 introduces the methodology for identifying correlation-driven market states and modeling their temporal transitions. Section 2.2 details the decision tree-based early warning model, including the design of interpretable state-informed features and projection-based labeling of new data. Section 3.2 presents empirical analysis of 15 years of CSI 300 data, covering market state classification, transition dynamics, and their link to crash periods. Section 4 evaluates the early warning model, focusing on feature configuration tuning, performance assessment, and interpretation of decision paths related to pre-crash signals. Section 5 concludes with key findings and implications for systemic risk monitoring and future research.

2 Methods

2.1 Market state classification and labeling

We conceptualize the financial market as a complex system, where its internal structure is reflected in the correlation patterns among individual stocks. A market state, in this context, is defined by the collective behavior of these correlations. To distinguish different states over time, we assess how the correlation structure evolves across different time windows, mainly following the approach proposed by Pharasi et al. [17, 19]. Specifically, we measure the similarity of correlation matrices computed over these windows, treating similar correlation patterns as indicators of the same underlying market state. To construct the correlation matrices, we begin with daily post-rights-adjusted prices of all stocks and assemble them into a data matrix. The logarithmic return of each stock i is calculated as

$$r_i(t) = \ln P_i(t) - \ln P_i(t-1),$$

where t denotes the end time point of the corresponding time window. Using these returns, we derive the correlation matrix $\rho(\tau)$ for each time window, with matrix elements given by the Pearson correlation coefficients between stock return pairs.

$$\rho_{ij}(\tau) = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\langle r_i^2 \rangle - \langle r_i \rangle^2}\sqrt{\langle r_j^2 \rangle - \langle r_j \rangle^2}}. \quad (1)$$

The average $\langle \ldots \rangle$ is evaluated within a fixed-length time window ending at time point τ, spanning M discrete time intervals. The subscripts $i, \quad j = 1, \ldots, N$ denote the stocks. To avoid ambiguity, we distinguish between two time-related indices: t refers to the actual trading day, while τ denotes the index of the rolling window, each of which spans a fixed number of trading days (e.g., $M = 22$). Thus, each window τ aggregates market information across the interval $[t - M + 1, t]$, enabling the construction of smoothed correlation matrices $C(\tau)$ for dynamic structural analysis. Since correlation matrices derived from short time series tend to be highly singular, we adopt the power mapping method to suppress noise in these matrices. This approach improves the stability and discriminative power of the correlation matrix spectra, as discussed in previous works like Guhr and Kälber [28]. The nonlinear exponentiation effectively reduces spurious correlations while preserving genuine structural relationships. In this approach, each Pearson correlation coefficient within a time window is transformed nonlinearly as

$$C_{ij}(\tau) = \operatorname{sign}\left(\rho_{ij}(\tau)\right) |\rho_{ij}(\tau)|^{1+\epsilon},$$

where $\epsilon \in (0, 1)$ is a parameter that controls the degree of noise suppression.

By applying this transformation, we obtain a sequence of noise-suppressed correlation matrices $C(\tau) = [C_{ij}(\tau)]_{N \times N}$ that capture the dynamic structure of stock return correlations over time. In our analysis, we use a rolling time window of $M = 22$ trading days (approximately 1 month), with a sliding step of $\Delta = 1$ day to ensure high temporal resolution.

We now define a similarity measure (or distance) between two correlation matrices, $C(\tau_1)$ and $C(\tau_2)$, evaluated at different ending time points τ_1 and τ_2, as:

$$d(\tau_1, \tau_2) \equiv \overline{|C(\tau_1) - C(\tau_2)|}, \tag{2}$$

where the overline denotes the average over all matrix elements. These pairwise distances are then used as input for Multidimensional Scaling (MDS), which projects the matrices into a low-dimensional space while preserving their mutual distances within an acceptable tolerance. MDS aims to map each correlation matrix $C(\tau)$ to a point in a lower-dimensional space, such that the pairwise distances between matrices are represented by a distance matrix $\zeta(\tau_1, \tau_2)$. This matrix is defined as:

$$\zeta(\tau_1, \tau_2) = \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,F_r} \\ \delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,F_r} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{F_r,1} & \delta_{F_r,2} & \cdots & \delta_{F_r,F_r} \end{pmatrix}, \tag{3}$$

where each entry $\delta_{i,j}$ represents the computed distance between matrices $C(\tau_i)$ and $C(\tau_j)$. The goal of MDS is to assign coordinates $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{F_r}$ in a low-dimensional Euclidean space such that the Euclidean distance $\|\mathbf{x}_i - \mathbf{x}_j\|$ approximates the original dissimilarities $\delta_{i,j}$, i.e.,

$$\|\mathbf{x}_i - \mathbf{x}_j\| \approx \delta_{i,j}, \quad \forall i, j \in \{1, 2, \ldots, F_r\},$$

where $\|\cdot\|$ denotes the Euclidean norm. In doing so, MDS provides a spatial representation of the F_r correlation matrices in a reduced-dimensional space $\mathbb{R}^D$.

```
Input: Set of F_r data points from MDS, number of clusters K
Output: Cluster assignments and centroids
Randomly initialize K cluster centroids among the F_r data points
repeat
Compute the Euclidean distance from each data point to all centroids
Assign each point to the cluster with the nearest centroid
Update each centroid by averaging the positions of all points in the cluster
until cluster assignments and centroid positions converge
```

Algorithm 1. *K*-Means Clustering for Market State Classification.

In this study, we select $D = 2$ or $D = 3$ to enable effective visualization. The resulting coordinates $\mathbf{x}_i$ are then used to generate two- or three-dimensional mappings that capture the structural relationships among correlation matrices over time. Structurally similar matrices are projected to nearby points, whereas dissimilar ones appear farther apart. Thus, MDS serves as a powerful tool to visualize and explore the dynamic evolution of market states through the similarity structure encoded in $\zeta(\tau_1, \tau_2)$.

We then apply K-means clustering to partition the F_r data points obtained from the MDS mapping into K distinct clusters, each defined by a cluster centroid. The algorithm proceeds through the following iterative steps:

To determine the optimal number of clusters K and the appropriate noise suppression parameter ϵ, we evaluate the clustering quality using the intra-cluster distance d_{intra}. This metric is defined as the average of the mean Euclidean distances from each point to its respective cluster centroid, aggregated across all clusters. For each candidate value of K, we perform the clustering under 1,000 different random initializations and compute the standard deviation of the resulting intra-cluster distances, denoted as $\text{std}_{\text{intra}}$. This procedure enables robust assessment of clustering stability and helps identify a reliable choice for K.

To determine the optimal number of clusters K, we introduce a pseudo-AIC criterion that balances intra-cluster dispersion and model complexity:

$$\text{AIC}_{\text{pseudo}} = 2K + n \ln\left(\text{std}_{\text{intra}}\right),$$

where K is the number of clusters, n is the number of samples, and $\text{std}_{\text{intra}}$ is the average intra-cluster standard deviation.

While this formulation resembles the structure of the Akaike Information Criterion (AIC), it is not derived from a traditional likelihood framework. Rather, it draws on a probabilistic interpretation of K-means under isotropic statistical assumptions and inspired by the information-theoretic model selection principle of balancing model complexity and goodness-of-fit [29]. To avoid confusion, we refer to it as a *pseudo*-AIC, and provide a detailed justification of this formulation in Supplementary Appendix A. The value of *pseudo*-AIC balances model fit and complexity, penalizing

Input: Daily closing prices P_t of a selected market index
Parameters: Time window length N, drawdown threshold η
Output: Labels $y_t \in \{0,1\}$ for each trading day
for each time window of length N (using sliding or fixed approach) **do**
for each trading day t in the window **do**
Compute drawdown: $\text{Drawdown}(t) = \frac{P_t}{\max(P_{\text{window}})}$
end for
if there exists t such that $\text{Drawdown}(t) < \eta$ and t is not the first day **then**
Let MaxDate be the date of peak price in the window
Mark the day before MaxDate as start of drawdown (Start_i), assign $y = 1$
Find the first day t' after Start_i such that $\text{Drawdown}(t') \geq \eta$
Mark the day before t' as end of drawdown (End_i)
Store drawdown window [Start_i,End_i]
end if
end for
Merge all intervals [Start_i,End_i] to form disjoint market crash periods
Assign $y_t = 0$ for all days t outside of these periods

Algorithm 2. Labeling Market Crash Cases from Market Price Data.

models with a large number of clusters. We select the optimal number of clusters K and the noise suppression parameter ϵ by minimizing the AIC value. This ensures that the selected model achieves both compact and stable intra-cluster structure while avoiding overfitting caused by an excessively large K.

In our analysis, we primarily consider cases where $K \geq 4$, as configurations with fewer than four clusters ($K \leq 3$) are deemed insufficient to capture the diversity of market states in a realistic manner. Once the optimal values of K and ϵ are determined, we compute the average correlation matrix within each cluster. These clusters are then sorted in ascending order based on their average correlation levels and assigned sequential labels corresponding to market states: $S1, S2, S3, \ldots$

2.2 Early warning model for market crash using machine learning methods

The first step in constructing the machine learning-based early warning model is to define the binary classification labels: market crash cases ($y = 1$) and non-crash (normal) cases ($y = 0$). These labels can be generated based on the historical daily closing prices of a representative market index. The labeling algorithm proceeds as follows:

In this paper, we employ the decision trees to construct the market crash early warning model. This modeling choice is motivated by several considerations. Decision trees are inherently simple yet highly interpretable, offering a clear hierarchical structure of decision rules that makes them particularly well-suited for applications requiring transparent reasoning. Second, they are relatively robust to overfitting, especially when the model complexity (e.g., tree depth) is properly controlled. Third, and most importantly in our context, decision trees are well-suited to scenarios with a limited number of input features. Since our early warning model is constructed based on market state transitions—resulting in a compact, semantically meaningful feature set—the ability of decision trees to perform reliable classification without requiring high-dimensional input is a critical advantage. These properties allow the model to maintain strong generalization performance while ensuring interpretability for financial analysts and policy researchers.

To enhance predictive performance, there are three types of features to be implemented to extract patterns associated with market state transitions in pre-crash dynamics:

- Proportions of Specific States in the Past L_1 Windows: $p_k^{L_1}$. Market states are first categorized into K distinct types. Over the most recent L_1 time windows, the frequency of each state is recorded and converted into proportions to reflect their relative prevalence. Let $n_{i,k}$ denote the number of times for market state k appears in the i-th window. The proportion of state k across these windows is computed as:

$$p_k^{L_1} = \frac{\sum_{i=1}^{L_1} n_{i,k}}{L_1} \quad (k = 1, 2, \ldots, K) \tag{4}$$

This process yields K features, represented by the vector $(p_1, p_2, \ldots, p_K)$, which captures the medium-term structural distribution of market states. These features are designed to reflect the prevailing level of market stability within the observed period.

- Ratio of Upward to Downward State Transitions in the Past L_2 Windows: r^{L_2}. An *upward transition* refers to a shift from a state with a lower index to one with a higher index (e.g., from S_3 to S_5), while a *downward transition* refers to the reverse (e.g., from S_4 to S_3). Let m_i and n_i denote the number of upward and downward transitions observed in the i-th window. The ratio over the most past L_2 windows is computed as:

$$r^{L_2} = \frac{\sum_{i=1}^{L_2} m_i}{\sum_{i=1}^{L_2} n_i}. \tag{5}$$

Although the state indices themselves are agnostic to specific economic meanings, empirical observations—both from our subsequent empirical results and related literature—suggest that transitions toward higher-indexed states tend to coincide with increased structural complexity or market disorder. Hence, a high value of r^{L_2} may indicate an ongoing shift in market conditions that precedes a turbulent phase. Conversely, a lower ratio may imply a trend toward greater structural coherence or reduced systemic stress. By capturing the relative directional trend in state transitions, this feature provides the model with a coarse-grained but informative indicator of evolving market stability.

(3) Evolutionary Path of Market States in the Nearest L_3 Windows: π^{L_3}. The sequence of market states over the most recent L_3 time windows is encoded as a categorical transition pattern. Let S_i denote the identified market state in the i-th window. The observed pattern is then represented as an ordered sequence:

$$\pi^{L_3} = \left(S_1 \to S_2 \to \cdots \to S_{L_3}\right). \tag{6}$$

Each unique sequence of L_3 consecutive states forms a categorical feature that reflects the short-term evolutionary path of market structure. Some simple examples include $\pi^3 = (S_2 \to S_3 \to S_2)$, $\pi^3 = (S_3 \to S_4 \to S_5)$, or $\pi^4 = (S_1 \to S_3 \to S_4 \to S_4)$. The total number of distinct features generated depends on both L_3 and the total number of state categories K, with longer sequences enabling more granular pattern differentiation.

These pattern-based features are particularly useful for capturing local dynamics, including sudden structural shifts, oscillatory behavior, or persistent regime patterns within stable/unstable conditions. By encoding the short-run evolutionary trajectory of market states, this feature set complements other indicators focused on medium- and long-term structural characteristics.

Given the asymmetric costs associated with classification outcomes in this context—where missing a true crash event (false negative) may result in significantly more severe consequences than issuing a false alarm (false positive)—the model selection process is designed to prioritize recall. Among models achieving the highest recall, further selection is made by jointly considering overall classification accuracy and minimizing the false positive rate (FPR), thereby ensuring both sensitivity and precision in practical application.

To address the issue of model adaptability over time, we further consider how to incorporate newly arriving data without compromising model consistency. A key challenge lies in assigning market state labels to new incoming time points. Re-clustering the entire dataset whenever new data becomes available is computationally expensive and introduces instability, particularly near cluster boundaries, where minor perturbations may lead to significant reassignments and undermine the validity of previously extracted features. Moreover, substantial updates in data may necessitate a recalibration of clustering hyper-parameters, including the optimal number of clusters and the noise suppression coefficient. To mitigate these issues, we propose the following procedure: compute the similarity (or distance) between the new data and the historical correlation matrices defined in Section 2.1, apply MDS algorithms to project the new observation into the same low-dimensional space, and subsequently assign its cluster label using a computationally efficient classifier (e.g., SVM) trained on the original MDS-embedded data.

3 The state transition behavior of the Chinese Stock Market

3.1 Data

This study utilizes the daily post-rights-adjusted prices of constituent stocks from the CSI 300 Index. The CSI 300, comprising 300 large-cap and highly liquid stocks from both the Shanghai and Shenzhen stock exchanges, represents approximately 60% of the total market capitalization of China's A-share market. As such, it serves as a key indicator for tracking the overall performance of the Chinese stock market. Its constituents are typically industry-leading enterprises that are highly responsive to sectoral and macroeconomic developments. Fluctuations in their stock prices tend to promptly reflect market sentiment and investor expectations, making them particularly suitable for investigating market state transitions.

Focusing on CSI 300 constituents allows the study to capture systemic market dynamics using a representative sample set, balancing analytical validity with computational efficiency. For the empirical analysis, there are 285 stocks selected from the CSI 300 Index, covering a 15-year period from January 2010 to December 2024, corresponding to $T = 3643$ trading days. To ensure the temporal consistency and comparability of correlation matrices $C_{ij}(\tau)$, we included only those stocks that were part of the CSI 300 on 4 January 2010, and remained continuously listed throughout the entire study period. The data is sourced from the RESSET database. All stock prices are post-rights-adjusted, a standard adjustment method that incorporates the effects of corporate actions including stock splits, cash and stock dividends, and rights offerings. This adjustment ensures that historical price series reflect true economic returns and are free from artificial jumps, thereby allowing for consistent return calculation across the full 15-year sample period.

3.2 Identification of market states

We compute the logarithmic returns of the daily post-adjusted closing prices for each stock and generate a sequence of 3,621 correlation matrices of dimension 285 using a sliding window approach, with the window length set to $M = 22$—approximately one trading month—and the sliding step $\Delta = 1$ to ensure sufficient granularity in state identification. To determine the optimal number of market states K, we employ the K-means clustering procedure as introduced in Algorithm 1 of Section 2.1, and apply the Akaike Information Criterion (AIC) to penalize excessive model complexity. Specifically, we examine how the number of clusters K and the noise-suppression parameter ϵ interact under varying initialization conditions, while tracking the stability of the clustering outcomes. The convergence condition is defined as:

- Convergence threshold: relative change in centroid position $< 10^{-5}$;
- Maximum iterations: 20 (to avoid oscillation in unstable cases).

Figure 1 presents the landscape of pseudo AIC-adjusted standard deviations across a grid of cluster counts ($K = 1$ to 10) and noise-suppression levels ($\epsilon = 0.0$ to 0.95). Color intensity reflects the magnitude of the standard deviation, with darker regions corresponding to more stable clustering outcomes. Each parameter configuration is evaluated over 1,000 independent K-means initializations, and the minimum observed pseudo AIC-adjusted standard deviation is interpreted as identifying the most robust and reliable clustering structure.



FIGURE 1
Stability landscape for clustering: Pseudo AIC-Adjusted standard deviation across K and ϵ.

As discussed in Section 2.1, although configurations with $K < 4$ sometimes yield lower intra-cluster standard deviations, we restrict our analysis to $K \geq 4$ to ensure sufficient interpretability and structural resolution. This choice is supported by prior studies in complex systems and financial market dynamics, which suggest that markets typically evolve through multiple qualitatively distinct regimes which are characterized by different levels of systemic co-movement and risk [11, 30]. Clustering configurations with fewer than four states often fail to capture this heterogeneity, resulting in overly coarse classifications that obscure meaningful structural transitions. For the CSI 300 Index, the configuration yielding the most stable and interpretable clustering outcome is ultimately determined to be $K = 5$ with $\epsilon = 0.0$.

Figure 2a presents a two-dimensional projection of the clustering result onto the xz-plane, derived from the optimal three-dimensional embedding obtained via MDS. This projection facilitates intuitive visualization of the cluster separation in reduced dimensions. The full three-dimensional representation of the MDS result, as shown in Figure 2b, provides a more comprehensive view of the geometric structure underlying the correlation matrices. Together, these visualizations demonstrate that the five identified market states are well-separated in the embedded space, suggesting the presence of latent structural patterns in market behavior over time.

Figure 3 presents example heatmaps of stock return correlation matrices for 285 constituent stocks, each selected from a representative trading day corresponding to one of the five identified market states under the configuration $\epsilon = 0.0$. A notable progression is observed across the states: as the market transitions upward (i.e., from lower- to higher-indexed states), the red intensity in the heatmaps increases, reflecting a rising level of average pairwise correlation among stocks. This pattern suggests a shift in market structure—from a relatively fragmented condition with weak inter-stock dependencies to a more cohesive regime marked by stronger co-movement. Such transitions often coincide with phases of elevated systemic dynamics, potentially signaling the emergence of market-wide trends or collective reactions to macro shocks.

To capture the temporal dynamics of market conditions, we examine the evolution of state occupancy probabilities over time. The color blocks in Figure 4 represent the relative frequency of each market state within semiannual windows. Specifically, the probability of a state in each window reflects the proportion of trading days classified under that state. Blue, orange, and green blocks correspond to lower-indexed market states, while red and purple denote higher-indexed states. Several noteworthy patterns emerge. Periods such as 2014, the second half of 2016 through 2017, and the second half of 2020–2021 are dominated by lower-indexed states, suggesting relatively stable market conditions. In

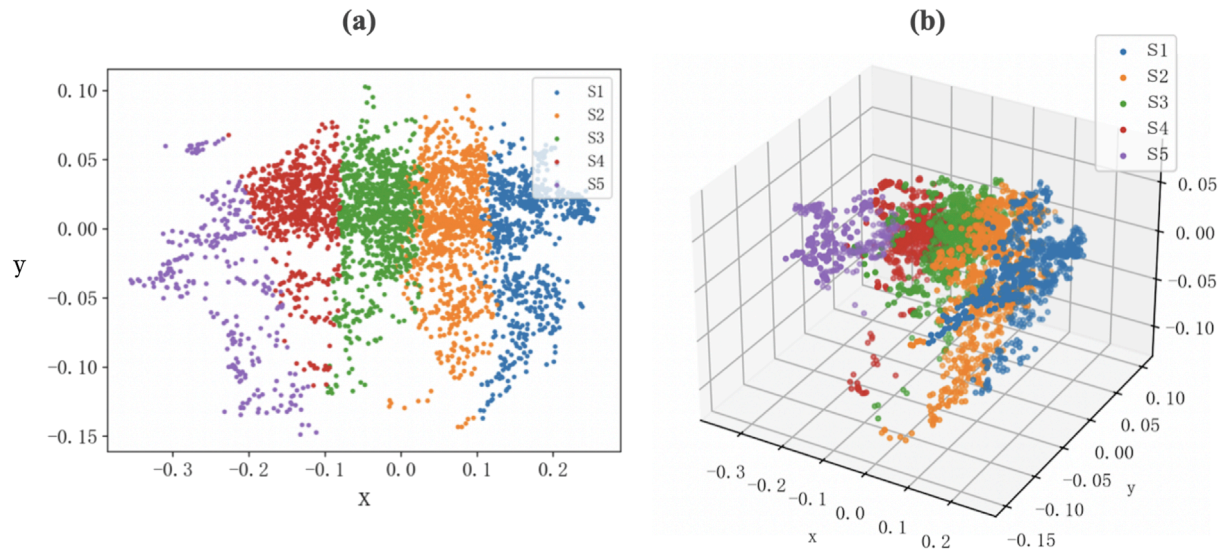


FIGURE 2
Multidimensional scaling (MDS) visualization of market state clustering.

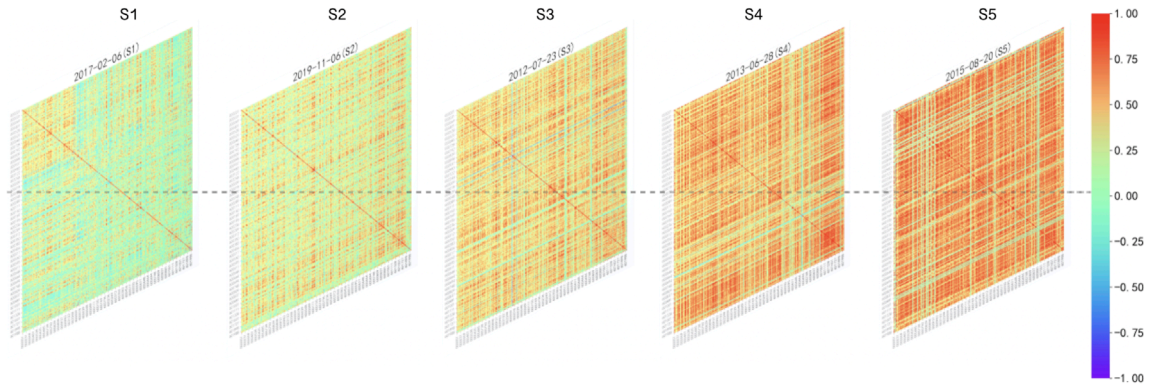


FIGURE 3
Representative correlation structures across market states.

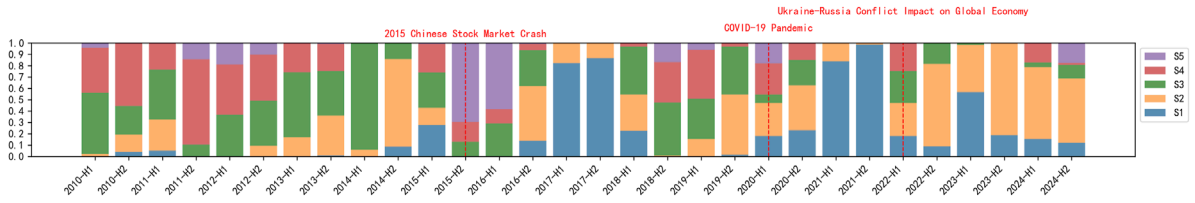


FIGURE 4
Semiannual market state landscape of Chinese stock market.

contrast, the window spanning 2015 to early 2016 is characterized by a pronounced transition to high-indexed states, aligning with the 2015 A-share market crash—one of the most severe episodes of turbulence in recent years. This temporal clustering of market states provides intuitive validation for the proposed state classification and its relevance to real-world market events.

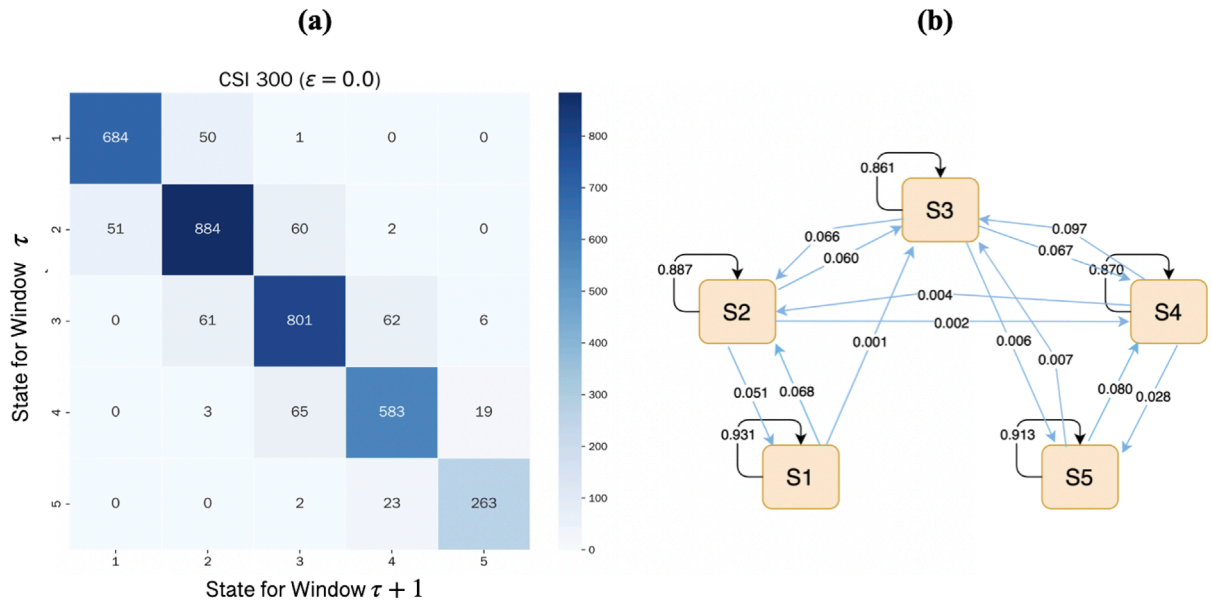


FIGURE 5
Empirical transition matrix and probabilities among market states.

3.3 Market states transition

To better understand the dynamic evolution of market conditions, we analyze the empirical transition behavior among the identified market states. Figure 5a presents the empirical state transition matrix of the CSI 300 Index, obtained under the optimal clustering setting $k = 5, \epsilon = 0.0$. The transition matrix exhibits a near-tridiagonal structure, suggesting that state changes tend to occur gradually—either remaining in the same state or shifting to adjacent states. This reflects the market's inertia and path-dependent nature in its structural evolution. Transitions that leap multiple levels—from lower-indexed states directly to a high-indexed state (e.g., from S_1 or S_2 to S_5)—are virtually nonexistent, highlighting the rarity of abrupt structural shifts without intermediate transitions. Instead, gradual escalations are more common. Notably, we observe a limited number of transitions from S_2 to S_4 (2 times), from S_3 to S_4 (62 times), from S_3 to S_5 (6 times), and from S_4 to S_5 (19 times), which may be interpreted as structural precursors to periods of elevated market stress.

Figure 5b visualizes the corresponding transition probabilities using a directed graph, where edge annotations indicate the estimated transition probabilities. Most of the mass is concentrated along self-loops and adjacent-state arrows, reinforcing the local stability of market states. Furthermore, the long-run stationary distribution derived from the estimated Markov chain closely aligns with the empirical frequency distribution:

$$P(S1, S2, S3, S4, S5) = (0.20, 0.28, 0.26, 0.18, 0.08),$$

which closely mirrors the empirical frequency of observed states. This consistency reinforces the robustness of the five-state market classification and suggests that the estimated state transitions capture meaningful and persistent structural dynamics in the market.

3.4 Relationship between market states and market crashes

To investigate whether market crashes tend to occur under specific structural states of the market, we examine the relationship between identified market states and empirically observed crash periods. The five market states of the CSI 300 Index are ordered by ascending average pairwise correlation, with the mean correlations given by $[\overline{C^{S1}(\tau)}, \overline{C^{S2}(\tau)}, \overline{C^{S3}(\tau)}, \overline{C^{S4}(\tau)}, \overline{C^{S5}(\tau)}] = [0.19, 0.30, 0.41, 0.52, 0.61]$, displaying a clear linear progression from weakly to strongly correlated market conditions.

Table 1 summarizes the start and end dates of crash episodes, the number of crash cases detected in each episode, the average correlation level during each crash period, and the corresponding market state assigned. These results are based on the proposed crash cases labeling procedure, i. e., Algorithm 2, with time window length $N = 10$ and a drawdown threshold of $\eta = 0.08$. To map each crash episode to a market state, we calculated the average correlation within the episode and classified it using the midpoint between adjacent state correlation means, i.e., $\frac{1}{2}(\overline{C^{S_i}(\tau)} + \overline{C^{S_{i+1}}(\tau)})$.

Based on Table 1, different crash episodes are mapped to different market states, forming an initial distribution of state assignments as $(S_1, S_2, S_3, S_4, S_5) = (0, 3, 4, 6, 9)$. However, some of these episodes are notably brief—lasting fewer than three time windows—and are likely driven by exogenous shocks rather than the result of endogenous market dynamics such as herding behavior or feedback amplification. To focus on more persistent and structurally significant crashes, we exclude these short-lived cases by removing periods with fewer than four cases. The adjusted distribution then becomes $(S_1, S_2, S_3, S_4, S_5) = (0, 1, 0, 4, 8)$, clearly indicating that most sustained crashes occur under higher-indexed states, particularly States 4 and 5. This result reinforces the association between elevated systemic co-movement and the

TABLE 1 Crash period characteristics and corresponding market states.

No. of crash periods	Start	End	No. of crash cases	Average correlation	State	Max drawdown
1	2010-04-14	2010-05-18	10	0.498	4	20.2%
2	2010-06-22	2010-07-05	3	0.467	4	9.7%
3	2010-11-08	2010-11-24	7	0.617	5	12.5%
4	2011-12-02	2011-12-15	1	0.329	2	8.5%
5	2013-05-28	2013-07-01	11	0.542	4	18.3%
6	2015-01-26	2015-02-06	1	0.383	3	8.2%
7	2015-06-08	2015-07-01	15	0.612	5	31.6%
8	2015-07-23	2015-08-06	7	0.664	5	10.3%
9	2015-08-10	2015-09-07	10	0.657	5	25.9%
10	2015-12-22	2016-02-01	18	0.714	5	26.4%
11	2018-01-26	2018-02-12	3	0.429	3	12.3%
12	2018-06-13	2018-06-28	2	0.422	3	9.6%
13	2018-07-24	2018-08-06	1	0.387	3	8.6%
14	2018-09-28	2018-10-18	5	0.624	5	11.5%
15	2019-04-19	2019-05-09	4	0.548	4	12.6%
16	2020-01-13	2020-02-05	3	0.714	5	12.3%
17	2020-03-05	2020-03-24	7	0.554	4	16.1%
18	2021-02-10	2021-03-10	5	0.288	2	14.4%
19	2022-03-01	2022-03-16	4	0.678	5	13.8%
20	2022-04-14	2022-04-26	2	0.542	4	9.7%
21	2022-10-18	2022-10-31	1	0.291	2	8.6%
22	2024-10-08	2024-10-17	4	0.693	5	11.0%

onset of significant market downturns. Figure 6 illustrates the time series of daily closing prices for the CSI 300 Index, with shaded areas indicating crash periods. Each color corresponds to the market state assigned to the crash episode: orange for S_2, green for S_3, red for S_4, and purple for S_5. As shown, these periods align well with major modification of the index, further validating the classification.

Overall, the results suggest a strong association between elevated-correlation market states and the occurrence of crashes. In particular, States 4 and 5 appear to signal heightened structural instability. This implies the Markov transition diagram of states in Figure 5b can be served to synchronously monitor the market's stability status. However, it is important to emphasize that such state classifications should be interpreted as contemporaneous reflections of market conditions, rather than forward-looking predictors. Whether they provide sufficient information for lead time warning is addressed in the following section.

4 Early warning of stock market crashes based on state transition information

This section explores the predictive value of market state transitions in generating early warnings for stock market crashes. Building on the classification framework established in previous sections, we aim to determine whether temporal patterns in market state dynamics can be effectively utilized to identify imminent periods of elevated market risk.

Based on the results in Table 1, we identify a total of 124 crash cases. For each of these, we label the trading day immediately

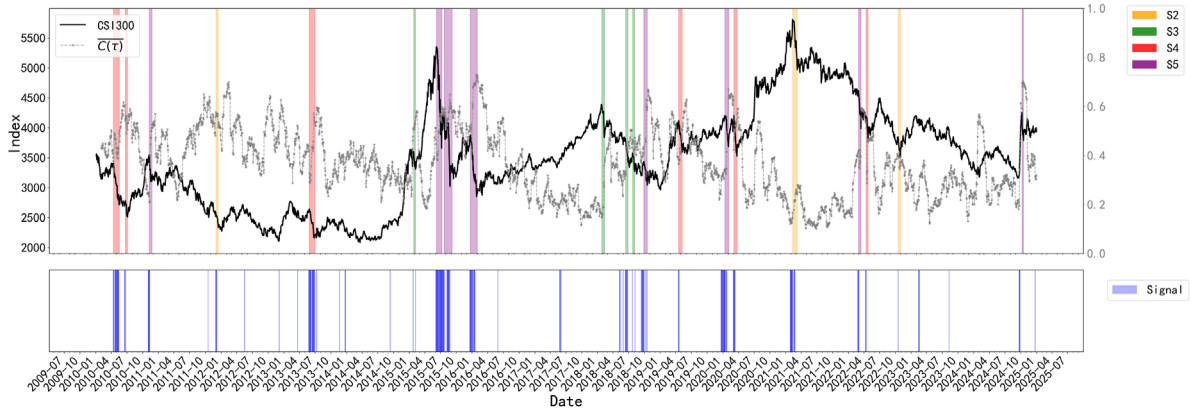


FIGURE 6
Crash periods and corresponding market states in the CSI 300 Index, along with the rolling forecasting performance of the early warning model. The top panel displays the CSI 300 Index (black) and the average correlation $\overline{C(\tau)}$ (gray dotted line), with vertical colored bands indicating identified crash periods and their dominant market states (S2–S5). The bottom panel shows real-time warning signals produced by the model under a rolling forecasting framework. Most signals are triggered shortly before or at the onset of crash periods, demonstrating the model's ability to provide timely early warnings.

TABLE 2 Top 10 best feature configurations.

Rank no.	L_1	L_2	L_3	Precisio	Recall	FPR
1	35	40	6	0.9636	0.9677	0.0359
2	35	40	5	0.9633	0.9677	0.0362
3	40	40	6	0.9630	0.9677	0.0365
4	35	40	4	0.9621	0.9677	0.0374
5	35	40	3	0.9615	0.9677	0.0380
6	40	40	5	0.9603	0.9677	0.0392
7	40	35	6	0.9587	0.9677	0.0410
8	40	40	4	0.9582	0.9677	0.0413
9	40	40	3	0.9574	0.9677	0.0423
10	40	30	4	0.9533	0.9677	0.0465

preceding the start of the case as $y = 1$, representing a positive (crash-imminent) sample. All other trading days outside the crash periods are labeled as $y = 0$. This procedure yields 124 positive samples and 3,289 negative samples. To mitigate the severe class imbalance in the dataset, the $y = 1$ samples are replicated 26 times, thereby constructing a balanced classification dataset suitable for supervised learning.

4.1 Model tuning and implication

To systematically examine how the temporal structure of state-based inputs affects the performance of crash early warning, we adopt the methodology outlined in Section 2.2, varying the window lengths used to construct the three categories of state-derived variables. Specifically, the candidate window lengths L_1 for the proportion of each market state $p_k^{L_1}$ and window lengths L_2 for the ratio of upward to downward state transitions r^{L_2} are both set to $(10, 15, 20, 25, 30, 35, 40)$. For the third feature type—capturing the recent evolutionary paths of market states—the window size L_3 is selected from $(3, 4, 5, 6)$. The smaller values of L_3 are chosen deliberately to reduce the combinatorial complexity of the resulting categorical pattern space, thereby enhancing model generalization and mitigating the risk of overfitting. These three temporal window features are designed to capture complementary structural and temporal dimensions of market dynamics. Specifically, L_1 encodes the medium-term proportion of time the market resides in each state, which reflects persistent structural tendencies; L_2 captures the directional momentum of market transitions—quantifying how frequently the market moves into more fragile or stable states—thus acting as a proxy for regime-level drift; and L_3 records short-term transition motifs, designed to detect recent local fluctuations that may precede regime shifts. The design draws inspiration from complex systems thinking, where multiscale memory effects and structural inertia often play a critical role in the lead-up to systemic events.

In total, 196 different (L_1, L_2, L_3) combinations are evaluated using a decision tree classifier with 5-fold cross-validation. Table 2 lists the top 10 configurations first ranked by recall and then ranked by precision. The best-performing configuration is $(L_1, L_2, L_3) = (35, 40, 6)$, which achieves a highest recall of 96.77%, along with the precision of 96.36%, and a low FPR of 3.59%. This setting is selected as the optimal feature configures for the model, balancing sensitivity to crash signals with robustness to false alarms. All performance metrics reported in this section are obtained via 5-fold cross-validation, where the dataset is randomly partitioned into five subsets. In each round, four subsets are used for training and one for validation, ensuring that every observation is tested exactly once. This repeated re-training procedure provides an estimate of model

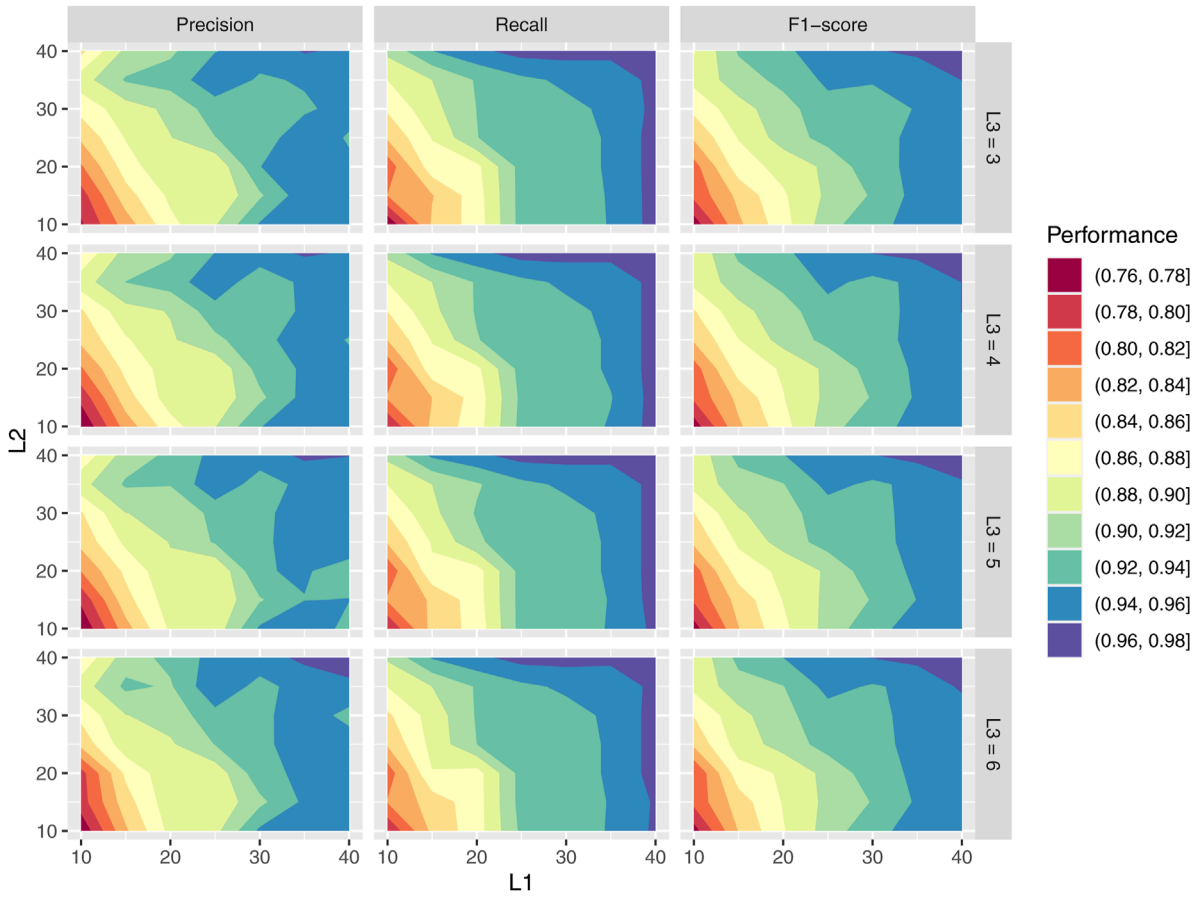


FIGURE 7
Heatmaps of precision, recall, and F1-Score with hyper-parameters combinations L_1,L_2,L_3.

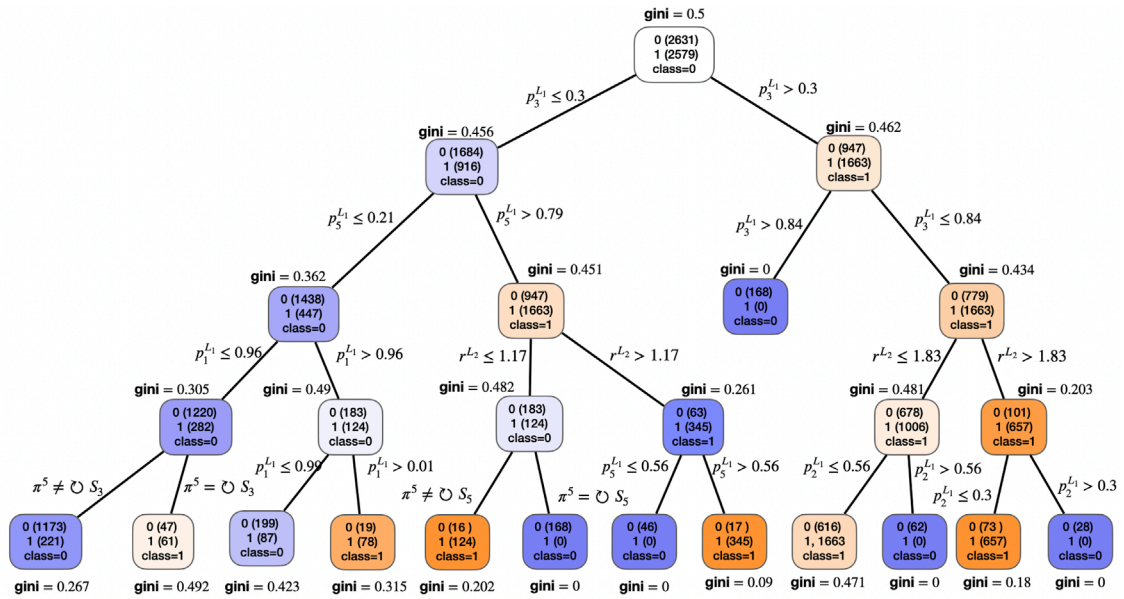


FIGURE 8
Decision tree structure under the optimal window configuration (L_1,L_2,L_3) = (35,40,6). Each node shows the predicted class, Gini coefficient, the number of samples labeled as crash ($y = 1$) and non-crash ($y = 0$), and the total number of samples in parentheses. Warm-colored nodes indicate classification into the crash category ($y = 1$), while cool-colored nodes correspond to non-crash classification ($y = 0$). The notation $\circlearrowleft S_i$ denotes persistent stay in state S_i, represented as a sequence such as ($S_i \to S_i \to S_i \to S_i \to S_i$).

TABLE 3 Comparison of early warning models using 5-fold cross-validation.

Model	Precision	Recall	F1-score	FPR
MST model	**0.964**	**0.968**	**0.966**	**0.037**
VIX-like model	0.863	0.755	0.803	0.120
Spectrum model	0.677	0.904	0.773	0.347
Absorption ratio	0.783	0.828	0.804	0.229

The maximum value in this column is indicated in bold.

generalization under moderate sample variability. To further assess the robustness of the selected feature configuration, we conducted an additional evaluation using repeated 5-fold cross-validation. Specifically, we repeated the 5-fold CV procedure 20 times with different random splits to estimate the variability of model performance. Based on these repetitions, we calculated confidence intervals for key performance metrics. The 95% confidence interval for precision is [96.2%, 96.5%], for recall is [96.7%, 96.8%], for F1-score is [96.5%, 96.7%], and for the false positive rate (FPR) is [3.4%, 3.7%]. These narrow intervals confirm the model's stable and reliable predictive capacity under different sample partitions.

We did not construct a separate test set for final model evaluation. Given the rarity and temporal clustering of crash events, a fixed test split may lead to high variance and unreliable evaluation. Instead, our focus is on examining whether temporal features of market state transitions offer informative signals for early crash detection. In future work, more rigorous backtesting approaches—such as Combinatorial Purged Cross-Validation (CPCV) [31]—could be explored to validate model performance in a production-like environment.

To provide a more detailed and intuitive assessment of how different window length combinations contribute to model performance, we visualize the precision, recall, and F1-score across all 196 (L_1, L_2, L_3) configurations in Figure 7. Each panel corresponds to a fixed value of $L_3 \in 3, 4, 5, 6$, while the horizontal and vertical axes represent values of L_1 and L_2, respectively. The color shading in each cell reflects the performance level for a specific metric, with darker shades indicating better results.

The heatmaps reveal several consistent patterns. First, larger values of L_1 and L_2 generally yield stronger performance across all metrics, especially when $L_3 = 6$, confirming the advantage of incorporating longer historical context for state proportions and transition ratios. Second, the F1-score—balancing precision and recall—tends to peak in the upper-right regions of the grids, where both L_1 and L_2 lie in the range of 30–40. This aligns well with the previously identified optimal configuration $(L_1, L_2, L_3) = (35, 40, 6)$, providing visual validation for the selected model parameters. Importantly, the combination $(L_1, L_2, L_3) = (35, 40, 6)$ integrates features over both short- and medium-term windows, capturing persistent structural shifts as well as abrupt transitions.

Figure 8 illustrates the structure of a representative decision tree trained under the optimal window configuration $(L_1, L_2, L_3) = (35, 40, 6)$, which displays the first four layers. Interpretation of the decision rules reveals that the most informative paths leading to a crash warning (i.e., terminal nodes classified as $y = 1$) tend to involve long-term patterns in state proportions and transition ratios—particularly $p_5^{L_1}$ and r^{L_2}—rather than short-term state transition sequences.

Notably, one of the clearest decision paths leading to a crash classification occurs when $0.3 < p_3^{L_1} \leq 0.84$, and $r^{L_2} > 1.83$, suggesting that elevated presence of mid-to-high index states combined with a strong net upward transition tendency serves as a robust early warning signal. In contrast, the specific forms of short-term state transition patterns—such as $\pi^5 = (S_i \to \cdots \to S_j)$—do not appear to provide significant predictive power on their own. However, when such patterns indicate that the market has not persistently remained in state S_5 (i.e., $\pi^5 \neq \circlearrowleft S_5$), and are combined with medium- to long-term indicators such as $p_k^{L_1}$ and r^{L_2}, the model still achieves strong early warning performance. Specifically, long-term state occupancy proportions $p_5^{L_1}$ act as proxies for structural stress accumulation—indicating prolonged market exposure to vulnerable regimes. The transition ratio r^{L_2} captures directional asymmetries in regime shifts, with high values signaling drift toward higher fragility states. The short-term motif π^5 provides further granularity by detecting whether the market remains stable in a specific state (e.g., $\circlearrowleft S_5$) or experiences disruptive shifts. These layered indicators together help uncover latent warning signals preceding systemic crashes, supporting the interpretability and theoretical consistency of the model.

To further examine the practical utility of the proposed early warning model, we apply the optimized classifier in a rolling forecasting mode. At each time step, the model uses only past information to predict whether a warning signal should be triggered for the next period. This simulates a realistic deployment setting where market operators make real-time decisions based on evolving structural dynamics. The lower panel of Figure 6 visualizes the resulting forecasted warning signals over the full sample period. Most warnings align closely with the start of crash periods identified by our drawdown-based labeling scheme, often appearing several days in advance. These results confirm that the model generalizes well to unseen data and is capable of generating timely early warnings in real-world settings.

To benchmark the performance of our proposed market state transition (MST) framework, we compare it with three representative baseline models of early warning signals that are widely used in the literature: (1) volatility-based indicators, (2) spectrum-based indicators and (3) absorption ratio.

For the volatility model, we construct a VIX-like indicator based on the implied volatility of the China ETF 50 options, often referred to as the "China fear index." This index serves as a functional proxy for market-perceived risk in the Chinese context. For the spectrum-based category, we follow the methodology of Chakraborti et al. [20] to compute the maximum eigenvalue of the correlation matrix. For the absorption ratio, we use the implementation proposed by Kritzman et al. [6], which quantifies the proportion of variance absorbed by a fixed number of principal components.

To enable a fair comparison, we construct decision-tree-based classifiers for each of these indicators by tuning thresholds to generate binary crash warnings. All models are evaluated using the same labeled crash dataset and assessed under 5-fold cross-validation. Table 3 presents the results. Our MST-based framework consistently outperforms the benchmark models across all metrics,

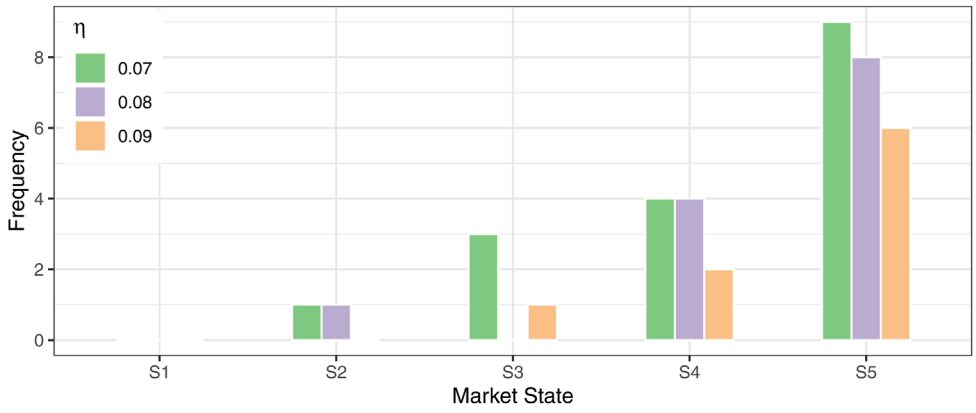


FIGURE 9
Sensitivity of crash-state distribution to drawdown threshold η. Frequencies of each market state observed during identified crash periods under three thresholds $\eta = 7\%, 8\%, 9\%$ are compared. Results confirm the consistency of crash-state associations.

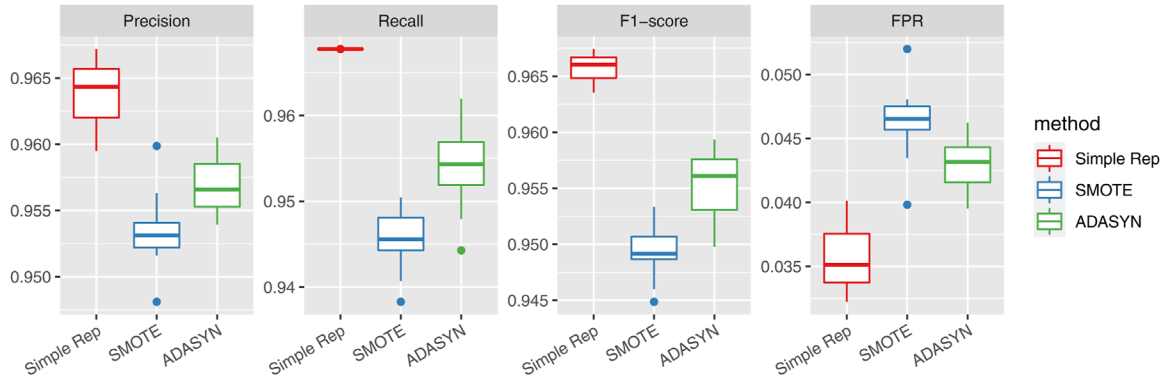


FIGURE 10
Model performance under different resampling methods. Bars represent precision, recall, F1-score, and false positive rate (FPR) for simple replication, SMOTE, and ADASYN. The results demonstrate that the model's performance remains stable across various class-balancing strategies.

achieving the highest precision (0.964), recall (0.968), and F1-score (0.966), along with the lowest false positive rate (3.7%). These results highlight the advantage of structurally grounded, transition-driven features in anticipating systemic stress.

4.2 Adaptability to new data

To evaluate the model's adaptability in real-time environments, we assess its ability to incorporate newly arriving trading data using the procedure described at the end of Section 2.2. Specifically, new correlation matrices are projected into the original low-dimensional MDS space, and their market state labels are assigned using a Support Vector Machine (SVM) trained on the initial clustering results.

Results from 5-fold cross-validation indicate that the SVM with a radial basis function (RBF) kernel achieves a classification accuracy of 98.20%, demonstrating that new trading days can be accurately labeled with minimal computational cost. This confirms the model's suitability for real-time deployment and continuous market monitoring under dynamic conditions.

4.3 Sensitivity analysis

To evaluate the robustness of our results with respect to key design choices, we conduct two sets of sensitivity analyses focusing on crash labeling and resampling methods.

First, we examine the impact of changing the drawdown threshold η used in crash event identification. Figure 9 illustrates the distribution of market states associated with crash periods under three threshold settings: $\eta = 7\%$, 8%, and 9%. The patterns remain highly consistent, confirming that the core findings reported in Section 3.4—especially the concentration of crash periods in states S_4 and S_5—are not sensitive to the exact choice of η.

Second, we assess the influence of different resampling techniques used to address class imbalance in model training. In addition to the simple replication method (Simple Rep) used in the main analysis, we consider two widely used alternatives: SMOTE and ADASYN. As shown in Figure 10, the decision tree model maintains stable performance across all three approaches. While SMOTE and ADASYN yield slightly lower precision, recall and F1-score, the overall differences are modest, supporting the validity of

our original choice. These results demonstrate that the predictive effectiveness of our framework is not materially dependent on the specific sampling method adopted.

5 Conclusion

Grounded in the complex systems perspective, this paper applies correlation-based structural analysis to identify and characterize market states in the Chinese stock market, with the CSI 300 Index as a representative system. Using multidimensional scaling and K-means clustering on rolling correlation matrices, we identify five distinct market states that reflect varying levels of systemic co-movement. Our analysis confirms that these states are not merely statistical artifacts but correspond to meaningful structural configurations of the market.

Empirical results reveal that market state transitions are predominantly local, occurring between adjacent states, while abrupt multi-level jumps are rare. States also exhibit high temporal persistence, underscoring the gradual and inertial nature of structural evolution in financial systems. Notably, high-indexed states—those with elevated average correlations—are strongly associated with historically observed crashes, particularly when the market resides in State 4 or State 5, where systemic co-movement is most pronounced. While these states offer valuable insight into ongoing market fragility, they function primarily as contemporaneous indicators rather than forward-looking predictors. Together, these findings support the use of market state classification as a structurally grounded, real-time measure of systemic risk.

Based on the identified state structure, an early warning model is constructed using decision trees trained on state-derived temporal features. These features capture medium-term structural distributions, directional trends in state transitions, and recent evolutionary patterns. The model demonstrates high recall and precision across multiple validation settings, indicating strong robustness in identifying pre-crash conditions. Furthermore, real-time adaptability is ensured through a projection-based labeling procedure, allowing new data to be efficiently embedded and classified, thus enabling continuous and responsive market surveillance.

Several avenues for future research emerge from this work. One direction is to extend the proposed framework to cross-market or multi-asset systems, such as incorporating global indices, fixed income instruments, or cryptocurrency markets, to examine whether state transitions exhibit comparable structural patterns across different financial domains. Another promising line lies in enriching the modeling of state transitions themselves—e.g., using recurrent neural networks to capture nonlinear memory effects. Furthermore, combining structural state dynamics with behavioral signals, news flows may enhance the predictive power and interpretability of crash warnings. Finally, linking market state trajectories with macro-financial policy variables could offer new insights into the interplay between systemic risk and regulatory interventions.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

W-YP: Data curation, Formal Analysis, Software, Visualization, Writing – original draft. LL: Conceptualization, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. We acknowledge financial support from the National Natural Science Foundations of China (Grant No. 71771086).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1647667/full#supplementary-material

References

1. Sornette D. Why stock markets crash: critical events. In: *Complex financial systems*. 2nd edn. Princeton, NJ: Princeton University Press (2017).

2. Mantegna RN, Stanley HE. *Introduction to econophysics: correlations and complexity in finance*. Cambridge University Press (2000).

3. Bouchaud J-P. The (unfortunate) complexity of the economy. *Phys World* (2009) 22:28–32. doi:10.1088/2058-7058/22/04/39

4. Hommes C. Complex evolutionary systems in behavioral finance. In: *Handbook of financial markets: dynamics and evolution*. Elsevier (2009). p. 217–76.

5. Münnix MC, Shimada T, Schäfer R, Leyvraz F, Seligman TH, Guhr T, et al. Identifying states of a financial market. *Scientific Rep* (2012) 2:644–6. doi:10.1038/srep00644

6. Kritzman M, Li Y, Page S, Rigobon R. Principal components as a measure of systemic risk. *The J Portfolio Management* (2011) 37:112–26. doi:10.3905/jpm.2011.37.4.112

7. Billio M, Getmansky M, Lo AW, Pelizzon L. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J Financial Econ* (2012) 104:535–59. doi:10.1016/j.jfineco.2011.12.010

8. Lin L, Sornette D. "speculative influence network" during financial bubbles: application to Chinese stock markets. *J Econ Interaction Coord* (2018) 13:385–431. doi:10.1007/s11403-016-0187-7

9. Lin L, Guo X. Identifying fragility for the stock market: perspective from the portfolio overlaps network. *J Int Financial Markets, Institutions Money* (2019) 62:132–51. doi:10.1016/j.intfin.2019.07.001

10. Sandhu R, Georgiou T, Tannenbaum A. Ricci curvature: an economic indicator for market fragility and systemic risk. *Sci Adv* (2016) 2:e1501495. doi:10.1126/sciadv.1501495

11. Samal A, Pharasi HK, Ramaiah SJ, Kannan H, Saucan E, Jost J, et al. Network geometry and market instability. *R Soc Open Sci* (2021) 8:rsos.201734. doi:10.1098/rsos.201734

12. Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, et al. Early-warning signals for critical transitions. *Nature* (2009) 461:53–9. doi:10.1038/nature08227

13. Hoel EP. When the map is better than the territory. *Entropy* (2017) 19:188. doi:10.3390/e19050188

14. Chen M, Li N, Zheng L, Huang D, Wu B. Dynamic correlation of market connectivity, risk spillover and abnormal volatility in stock price. *Physica A: Stat Mech Its Appl* (2022) 587:126506. doi:10.1016/j.physa.2021.126506

15. Wu B, Huang D, Chen M. Estimating contagion mechanism in global equity market with time-zone effect. *Financial Management* (2023) 52:543–72. doi:10.1111/fima.12430

16. Yu D, Huang D. Cross-sectional uncertainty and expected stock returns. *J Empirical Finance* (2023) 72:321–40. doi:10.1016/j.jempfin.2023.04.001

17. Pharasi HK, Sharma K, Chatterjee R, Chakraborti A, Leyvraz F, Seligman TH Identifying long-term precursors of financial market crashes using correlation patterns. *Physica A: Stat Mech its Appl* (2018) 516:603–16.

18. Heckens A, Guhr T. A new attempt to identify long-term precursors for financial crisis in the market correlation structures. *arXiv preprint* (2021).

19. Pharasi HK, Sharma K, Chakraborti A. Dynamics of market states and risk assessment. *Physica A: Stat Mech its Appl* (2024) 631:129191.

20. Chakraborti A, Sharma K, Pharasi HK. Phase separation and scaling in correlation structures of financial markets. *J Phys Complexity* (2020) 2:015002. doi:10.1088/2632-072x/abbed1

21. Chetalova D, Schäfer R, Guhr T. Zooming into market states. *J Stat Mech Theor Exp* (2015) 2015:P01029. doi:10.1088/1742-5468/2015/01/p01029

22. Spelta A, Flori A, Pecora N, Buldyrev S, Pammolli F. A behavioral approach to instability pathways in financial markets. *Nat Commun* (2020) 11:1707–9. doi:10.1038/s41467-020-15356-z

23. Begušić S, Kostanjčar Z, Kovač D, Stanley H, Podobnik B. Information feedback in temporal networks as a predictor of market crashes. *Complexity* (2018) 2018:1–13. doi:10.1155/2018/2834680

24. Wang X, Zhao L, Zhang N, Feng L, Lin H. Stability of China's stock market: measure and forecast by ricci curvature on network. *Complexity* (2023) 2023:1–12. doi:10.1155/2023/2361405

25. Xing K, Yang X. How to detect crashes before they burst: evidence from Chinese stock market. *Physica A: Stat Mech Its Appl* (2019) 528:121392. doi:10.1016/j.physa.2019.121392

26. Cerqueti R, Gatfaoui H, Rotundo G. Resilience for financial networks under a multivariate garch model of stock index returns with multiple regimes. *Ann Operations Res* (2024). doi:10.1007/s10479-023-05756-x

27. Lim B, Zohren S, Roberts S. Detecting changes in asset co-movement using the autoencoder reconstruction ratio. *arXiv preprint arXiv:2002.02008* (2020).

28. Guhr T, Kälber B. A new method to estimate the noise in financial correlation matrices. *J Phys A: Math Gen* (2003) 36:3009–32. doi:10.1088/0305-4470/36/12/310

29. Burnham KP, Anderson DR. *Model selection and multimodel inference: a practical information-theoretic approach*. 2nd edn. Springer (2004).

30. Chakraborti A, Sharma K, Pharasi HK. Characterization of catastrophic instabilities: market crashes as paradigm. *arXiv preprint* (2018).

31. López de Prado M. *Advances in financial machine learning*. Wiley (2018).

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 07 August 2025
DOI 10.3389/fphy.2025.1614995

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Can-Zhong Yao,
South China University of Technology, China
Naixi Chen,
Donghua University, China
Pristanto Silalahi,
Atma Jaya University, Yogyakarta, Indonesia

*CORRESPONDENCE
Shanshan Jiang,
jss@nuist.edu.cn

RECEIVED 20 April 2025
ACCEPTED 15 July 2025
PUBLISHED 07 August 2025

CITATION
Chen S, Jiang S, Zhuang Y and Meng L (2025) Research on the risk contagion of banks holding subordinated debt from the perspective of risk preferences.
Front. Phys. 13:1614995.
doi: 10.3389/fphy.2025.1614995

COPYRIGHT
© 2025 Chen, Jiang, Zhuang and Meng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Research on the risk contagion of banks holding subordinated debt from the perspective of risk preferences

Shuting Chen[1], Shanshan Jiang[1]*, Yikang Zhuang[1] and Lingyi Meng[2]

[1]School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China, [2]School of Politics, Economics and International Relations, University of Reading, Reading, United Kingdom

Introduction: Under global integration, banking system stability is paramount for financial market stability. While subordinated debt enhances bank capital stability, it poses capital loss risks to creditor banks upon issuer bankruptcy. This paper examines the transmission mechanism of systemic risk arising from banks' mutual holdings of subordinated debt, specifically from a risk preference perspective.

Methods: We construct a multi-channel risk contagion model that comprehensively incorporates the effects of interbank lending, investment coupling, and mutual holdings of subordinated debt, explicitly integrating bank risk preference. The model's dynamics and systemic risk implications are analyzed through simulation.

Results: Simulation analysis reveals that bank systemic risk is significantly influenced by savings rates and their volatility, reserve requirements, and investment returns and their volatility. In contrast, the impact of interbank lending rates is found to be relatively small.

Discussion: These findings provide crucial theoretical support for regulatory agencies in identifying key risk contagion pathways and formulating effective preventive strategies. Furthermore, the study outlines important future research directions, including calibrating model parameters with real-world data and introducing more heterogeneous bank behavior assumptions.

KEYWORDS

systemic risks, subordinated debt, risk preference, banking network, dynamic evolution

1 Introduction

As the global economy continues to become more integrated, national financial markets are becoming increasingly interconnected. As the nucleus of the financial system, the banking system's business activities penetrate deeply at all levels, and once systemic risk occurs, it will not only trigger the crisis of the bank itself, but also quickly transmit it to the entire financial market like dominoes. The Bank for International Settlements defines systemic risk in financial markets as the risk that a substantial part of the financial system will be unable to fulfil its function of providing credit and collapse [1]. This

paper defines systemic risk based on the perspective of bank failure risk. Risk contagion is a key issue in the financial system. The theory of risk contagion suggests that the interconnectedness of financial institutions can lead to the rapid spread of distress in one institution to others, thereby triggering systemic risk. Risk contagion in this paper refers to the process by which the risk of one bank is transmitted to other banks through channels such as interbank lending and investment coupling, resulting in other banks also facing risk.

Current research on systemic risk in the banking system primarily focuses on analysis from the perspective of the interbank market. One category of research emphasizes the construction of financial network models to analyze the risk contagion resulting from direct credit relationships between banks [2–7]. For example, Qi et al. [4] used network topology theory to analyze the transmission mechanisms of credit shocks and liquidity shocks, as well as their impact on risk propagation within the banking network. The study found that liquidity shocks have a greater impact than credit shocks; Liu [5] analyzed the interactions between financial institutions and interbank lending, as well as the effects of leverage and market liquidity. The results emphasized the importance of understanding and controlling risk channels for maintaining banking stability; Macchiati [6] constructed an interbank market model using European banks' balance sheet information to assess systemic liquidity risk; Yanquen and Livan [7] found that credit provided by banks to the real sector exhibits a certain degree of overlap, thereby creating systemic risk. Another category of research focuses on the dynamic evolution of interbank lending activities and their dynamic impact on systemic risk [8–15]. For example, Capponi et al. [8] developed a dynamic model of interbank lending activities to reflect the dynamic changes in systemic risk; Ding et al. [10] used a multi-agent approach to model the interbank lending network, conducting an in-depth study of the risk contagion process among banks; Fan and Sheng [11] constructed a dynamic evolution model for assessing systemic risk in the banking sector under dual risk exposure, with results indicating that the evolution of contagion risk exceeds that of credit risk; Jin et al. [13]conducted stress tests on China's interbank lending network to study the risk contagion capacity of financial networks; Krause and Giansante [14] established an interbank lending network and found that the initial scale of bank failures is the dominant factor in whether contagion occurs.

The extant studies primarily analyze systemic risk from the interbank lending perspective, and some scholars also consider banks' common asset holdings [16, 17] and investment-related behaviors [18–21]. Wang and Li [16] found that the redistribution of assets within the two networks, namely, interbank loans and mutual asset holdings, significantly reduced the systemic risk; Zou et al. [19] developed a two-tiered network model of China's financial system, which included the interbank lending network and the cross-stockholding network. Their findings indicated that the contagion losses within the two-tiered network exceeded the total losses from the independent interbank credit network and the cross-stockholding network. Gao and Fan [20] introduced risk appetite coefficient and systemic shocks to construct a quantitative portfolio strategy model. Nguyen [22] found that subordinated debt mitigates bank risk-taking; Cappeni and Weber [23] studied the portfolio selection problem of banks through the losses caused by the spillover effect of price reductions; Shen and Li [24] found that risk contagion was also significant due to the combined effect of counterparty risk and common asset holding risk when there was no overlap in portfolios. Weber and Weske [25] examined the individual and combined effects of interbank liabilities, bankruptcy costs, markups, and cross-shareholdings on systemic risk; Wang et al. [26] found that the risk of contagion caused by excessive similarity of investment assets is the main cause of systemic risk. Jiang and Fan [27] considered the effects of interbank loans, portfolio overlap, and subordinated debt mutual ownership on bank stability; Yan et al. [28] established a multi-tier banking network structure with three channels: interbank loans, mutual ownership, and common asset holding.

The aforementioned models consider the impact of interbank lending, jointly held assets, or investment-related behaviors on systemic risk in the banking system, but most studies are based on single-channel or dual-channel approaches. Jiang and Fan [27] were the first to propose a multi-channel model that integrates interbank loans, overlapping investment portfolios, and cross-holdings of subordinated debt. However, their method assumes that the interest rates on all subordinated debt issued by banks are fixed, and the allocation process of debt holdings is random. This overlooks two key realities in financial markets: (1) Subordinated debt interest rates are inherently risk-sensitive, reflecting the financial health of the issuer; (2) Banks' investment choices (including debt holdings) are strategically driven by their risk preferences. To bridge this gap, our model introduces a risk-preference-based dynamic mechanism: (1) Debt issuance rates are differentiated based on the relative equity strength of the issuer; (2) Debt holdings are not random; according to risk premium theory, banks with lower risk preferences (risk-averse) tend to hold lower-yielding (safer) debt, while banks with higher risk preferences (risk-seeking) pursue higher-yielding (riskier) debt.

2 Model

2.1 Bank market structure with mutual holding of subordinated debt

In the banking system, bank bankruptcy is frequently due to insufficient liquidity and a bank's liquidity is closely correlated with its savings, investment, as well as other behaviors [29]. To examine the specific factors affecting bank systemic risk, this paper establishes a multi-channel risk transmission model on the basis of interbank holdings of subordinated debt with risk preferences, as illustrated in Figure 1. This model comprehensively considers the effects of interbank loans, investment in ordinary assets, and mutual holdings of subordinated debt on bank systemic risk. Different from previous studies, this paper further considers the heterogeneity of banks, setting different risk preference parameters for each bank to reflect the bank's preference for risk. Capital structure theory states that the cost of financing a firm is closely related to its capital structure. A bank's owner's equity is an important part of its capital structure, and higher owner's equity means that the bank's financial position is more robust and better able to absorb potential losses, thus reducing the cost of its debt financing. Therefore, this paper sets different subordinated debt interest rates based on the

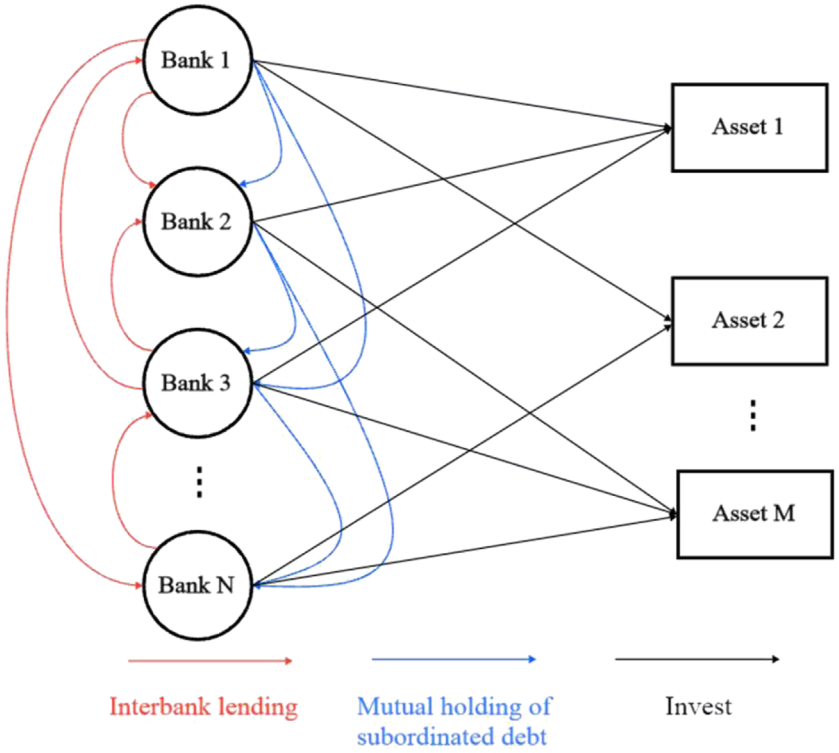


FIGURE 1
Interbank market structure.

ratio of each bank's owner's equity to the average owner's equity of all banks. At the same time, Subordinated debt is selected for holding based on the size of the risk preference parameters; banks with smaller risk preference parameters (i.e., more risk-averse) hold subordinated debt with smaller interest rates, while the opposite holds subordinated debt with larger interest rates.

In the actual banking system, banks usually hold subordinated debt with each other as a long-term capital replenishment mechanism to meet capital adequacy requirements. Subordinated debt can be traded in the secondary market, while interbank lending is mainly used to meet the short-term funding needs of banks and serves to manage liquidity, and is generally not transferred. In addition, the interrelation of interbank investment assets plays a critical part in the banking system's stability. Multiple banks may invest in the same asset, and once one bank goes bankrupt and conducts an asset liquidation, the price of the asset will fluctuate, affecting the other banks holding the asset. In addition to causing asset price volatility, the failure of a bank may result in the bank's inability to meet its loan and subordinated debt obligations, causing losses to the assets of creditor banks. In severe cases, it may even trigger a domino effect of consecutive bank failures. Insufficient liquidity of banks, which is closely correlated with factors such as investment, savings and lending, is the major factor of instability in the banking system. This article will use the model to analyze how these factors affect systemic risks in the banking system.

2.2 Model construction of multi-channel risk transmission among entities holding subordinated debt

In the actual banking system, the relationships between banks are complex and intertwined, involving not only direct relationships arising from interbank lending and holding subordinated debt, but also indirect relationships arising from joint investments in certain assets. In a random banking network, the interbank lending and the issuance and holding of subordinated debt are a process of random allocation. When banks have insufficient liquidity, they can borrow from other banks, with the link matrix represented by J, where J_{ij} takes a value of 0 or 1. $J_{ij} = 1$ indicates a possible lending relationship between bank i and bank j, while $J_{ij} = 0$ indicates no credit relationship between bank i and bank j. S represents the interconnection matrix of banks holding subordinated debt mutually, where when $S_{ij} > 0$, the quantity of subordinated debt that bank i holds in bank j is S_{ij}, and when $S_{ij} < 0$, the quantity of subordinated debt that bank j holds in bank i is S_{ij}, thus $S_{ij} = -S_{ij}$. The connection probability between banks is represented by P, where $0 \le P_{ij} \le 1$.

Under the random banking network scenario, it is assumed that at time t, there exist N_t banks, where N_t is a bounded integer. As the discrete time evolves through $t = 1, 2, 3, \ldots, T$, these N_t banks form direct and indirect connections. At time t, bank i's liquidity

$L_i(t)$ is given by Equation 1:

$$\boldsymbol{L_i(t) = \overline{L_i}(t) - D_i(t) - I_i(t) + \sum_{j=1}^{N_t} c_{ij}(t) B_{ij}(t)} \tag{1}$$

where $\overline{L_i}(t)$ is the liquid assets of bank i that have not been invested, distributed and dismantled at time t; $D_i(t)$ is the distribution of dividends of bank i at time t; $I_i(t)$ is the investment of bank i at time t; $c_{ij}(t)$ represents the interbank loan association of bank i with bank j at time t. When there is a loan association of bank i with bank j at time t, $c_{ij} = 1$, otherwise $c_{ij} = 0$; $B_{ij}(t)$ refers to the loan volume between bank i and bank j. When $B_{ij}(t) > 0$, bank i borrows money from bank j, and *vice versa*, bank i lends money to bank j. In stochastic banking networks, borrowing only occurs when banks are illiquid, which is consistent with how banks actually operate.

At each time, the bank will carry out savings, dividends and investment operations, these operations will make the bank's liquidity to alter, at time t, the bank's liquidity without investment, dividends and savings $\overline{L_i}(t)$ can be expressed by Equation 2:

$$\boldsymbol{\overline{L_i}(t) = L_i(t-1) + A_i(t) - A_i(t-1) - r_a A_i(t-1) + U_i(t)} \tag{2}$$

where r_a is the bank deposit interest rate, $A_i(t)$ represents the amount of deposit owned by bank i at time t, and its concrete expression is $A_i(t) = (1 + \sigma_A \eta_t)\overline{A}$, where σ_A is the standard deviation of the total bank deposits, $\overline{A}$ is the average value of total bank deposits, $\eta_t \in N(0,1)$. $U_i(t)$ denotes the investment profit earned in bank i at time t, which is closely related to the total investment amount of bank i at time t. Under the condition that bank i has not invested at time t, the total investment amount of bank i is expressed by Equation 3:

$$\boldsymbol{Y_i(t) = \sum_{j=1}^{M} Q_{ij}(t-1) g_j(t)} \tag{3}$$

where $Q_{ij}(t-1)$ is the amount of equity in asset j that bank i holds at time $t-1$, which is a dynamic value that reflects the realization of past investments and the development of new investments by bank i. $g_j(t)$ is the value of asset j at time t, and its expression is given by Equation 4:

$$\boldsymbol{g_j(t) = g_j(t-1)\left(1 + \delta_j(t)\right)} \tag{4}$$

where $\delta_j(t)$ is the yield of asset j at time t, and it is not a fixed value, but a dynamic value influenced by many factors such as macroeconomics and policies. In this paper, $\delta_j(t)$ obeys a normal distribution with mean τ, i.e., $\delta_j(t) \in N(\tau, \theta)$. Where τ is the mean yield of all investments and θ is the asset value volatility.

When the bank's liquidity is inadequate, it will choose to liquidate some assets to make up for the liquidity. At time t, the quantity of assets realized by bank i is $U_i(t) = q \times Y_i(t)$, where q is the ratio of realized investment. The aggregate investment assets owned by bank i after the investment is realized and before continuing the investment are $\overline{Y_i}(t) = \sum_{j=1}^{M} \overline{Q_{ij}}(t) g_j(t)$, where $\overline{Q_{ij}}(t)$ is the amount of shares in asset j owned by bank i following the realization of the partial investment. So the bank owner's equity before dividends and investments can be expressed as Equation 5:

$$\boldsymbol{\overline{V_i}(t) = \overline{L_i}(t) + \overline{Y_i}(t) - A_i(t) - (1 - r_b) B_i(t-1)} \tag{5}$$

where $B_i(t-1) = \sum_{j=1}^{N_{t-1}} B_{ij}(t-1)$ is the aggregate interbank debt owned by bank i at time $t-1$, r_b is the interbank offered rate. $(1 + r_b)B_i(t-1)$ denotes the sum of the interbank borrowing capital and interest that bank i has to repay at time t. When the bank generates investment income, it can pay dividends. The dividend of bank i at time t is determined by Equation 6:

$$\boldsymbol{D_i(t) = max\left\{0, min\left[U_i(t) - r_a A_i(t-1), \overline{L_i}(t) - \beta A_i(t), \overline{L_i}(t) + \overline{Y_i}(t) - (1+\chi) A_i(t)\right]\right\}} \tag{6}$$

where β is the deposit reserve ratio and χ is the capital saving ratio. Dividend distribution by banks should be based on three premises: $\overline{V_i}(t)/A_i(t) \geq \chi$, net yield greater than 0, and meeting the requirements of deposit reserve. After paying dividends, the bank needs to reinvest the residual cash to obtain more returns. The investment made by bank i at time t is expressed as Equation 7:

$$\boldsymbol{I_i(t) = min\left\{max\left[0, \overline{L_i}(t) - \beta A_i(t) - D_i(t)\right], \omega_i(t)\right\}} \tag{7}$$

where $\omega_i(t) = \overline{\omega}(1 + \sigma_\omega \mu)$ is the investment opportunity of bank i at time t, where $\overline{\omega}$ Is the average of all bank investment opportunities, σ_ω is the standard deviation of all bank investment opportunities, and μ obeys a normal distribution of $N(0,1)$. At time t, the amount of asset j invested by bank i is $\Delta Q_{ij}(t)$, which satisfies $I_i(t) = \sum_{j=1}^{M} \Delta Q_{ij}(t) g_j(t)$. After investment $Y_i(t)$ satisfies the formula $Y_i(t) = \sum_{j=1}^{M} Q_{ij}(t) g_j(t)$, where $Q_{ij}(t) = \overline{Q_{ij}}(t) + \Delta Q_{ij}(t)$.

2.3 Multichannel risk contagion model analysis of mutual subordinated debt holdings based on risk appetite

Within the framework of the banking system's network, if a bank holds liquid funds after completing its dividend distribution and investment activities, it is considered to be a potential lender of funds, capable of providing financing support to other banks. When a bank is illiquid, it will need to incorporate funds from the interbank market to continue operating as usual. In case of the indebted bank, when the funds received from other banks are able to cover the preceding debt, the repayment operation will be carried out, and at the same time, the liquidity of this bank will be reduced to nil. If the bank is unable to raise sufficient funds to repay its debts and interest on deposits, it will have to dispose of its assets until the funds from the sale of assets can cover its debts. If the bank is unable to pay its debts after selling its assets, it will face bankruptcy and start the process of liquidation of assets. The proceeds of the liquidation will be applied first to pay off deposits and the remainder will be repaid to creditor banks on a pro rata basis.

During the process of issuing subordinated debt, the interest rate r_i of subordinated debt of bank i is determined by the ratio of bank i's ownership interest to the average ownership interest of all banks. Its pricing follows the risk-return matching principle: banks with low capital adequacy ratios need to pay higher risk premiums to compensate investors. According to the Merton theory model, when equity declines, the risk of default increases exponentially, so the specific expression is as Equation 8:

$$\boldsymbol{f(r_i) = \overline{r} + \frac{e^{1 - \frac{v_i}{\overline{v}}} - 1}{1000}} \tag{8}$$

where $\bar{r}$ is the average subordinated interest rate of all banks, V_i is bank i's ownership interest, $\overline{V}$ is the average ownership interest of all banks. In the actual financial market, the interest rate at which a bank issues subordinated debt is affected by its own financial position. In assessing the credit rating of banks' subordinated debt, credit rating agencies will take the banks' financial indicators such as owner's equity and capital adequacy ratio as important references. Usually, the subordinated debt issued by banks with better financial conditions can obtain a higher credit rating, thus attracting investors with lower interest rates. This is consistent with the model's assumption that subordinated debt rates are negatively correlated with bank ownership equity. In the allocation process of subordinated debt, Bank i's risk preference parameter is ρ, reflecting the bank's level of risk preference. In the actual banking system, a bank's risk preference is influenced by a variety of factors and cannot be fully captured by simple deterministic rules. Therefore, in this paper, the risk preference parameter is a random number between 0 and 1. The larger the risk preference parameter of banks, the more they like risk, and they tend to hold subordinated debt with higher interest rate; otherwise, they are more risk-averse and tend to hold subordinated debt with lower interest rate.

In the dynamic development of the banking system, selling assets tends to cause asset prices to fall. Since a number of financial intermediaries invest in the same asset at the simultaneous time, financial intermediaries are indirectly connected to each other through the assets in which they invest. Movements in asset prices can influence multiple institutions holding the asset at the same time, and changes in the assets of one financial institution can have an effect on other financial institutions through the assets jointly held. When banks are illiquid and unable to access loans, they sell assets to cover the liquidity gap. Furthermore, when a bank fails, the failing bank will liquidate its portfolio of assets, and the assets sold will decline in value. Banks holding the identical assets will be hit by the devaluation of their assets, leading to a loss of equity, which in turn may lead to insolvency, triggering bank bankruptcy, which can further damage the assets of creditor banks. Through this evolutionary process, the initial shock of the interbank network continues to propagate through the system. When a bank sells an asset due to illiquidity, or a bank goes into liquidation due to bankruptcy, the asset is sold at a depreciated value. Here, a market effect function is modeled through the price impact function in Equation 9:

$$f\left(x_j^t\right) = e^{-\alpha x_j^t} \tag{9}$$

where x_j^t is the liquidation portion of asset j at time t. α represents the asset price sensitivity, i.e., the degree of price fluctuation resulting from the sale of the asset. In this paper, we take the value of α as 1.0536 in reference to the study of Caccioli [30], which indicates that when 10% of the assets are sold, the price of the assets will also fall by 10%, which can be approximated as the linear effect of asset sales on asset prices. Thus, the value of asset j at time t is as shown in Equation 10:

$$g_j(t) = g_j(t) f\left(x_j^t\right) \tag{10}$$

3 Model analysis

In the banking system evolutionary process, banks receive various ownership interests, liquid assets, dividends, *etc.*, depending on their business conditions and strategies, and some banks go bankrupt due to poor business strategies. Due to systemic risk, the failure of one or a few banks may trigger a domino effect of more consecutive bank failures. In order to more clearly characterize systemic risk, we define systemic risk as in Equation 11:

$$\boldsymbol{Risk}(t) = \frac{1}{T_S N_e} \sum_{i=1}^{N_e} \sum_{j=t+1}^{t+T} \frac{M_j^i}{N_j^i} \tag{11}$$

where T_S is the degree of time, $T_S = 200$ in this paper, N_e is the count of simulation repetitions, N_j^i is the count of banks that are not insolvent at moment j, M_j^i is the count of insolvent banks at the i th simulation.

In this paper, the largest simulation time step is 2000, 200 banks are used in the simulation, and the simulation is performed 10 times, which can adequately represent the dynamic features of the banking system. At each time step, the bank can randomly recover 35% of its assets. There are 150 assets in which the bank can invest. The initial equity capital of the owner follows a normal curve with a mean of 200. The deposit interest rate is less than the investment return rate, i.e., $r_a < \delta$, which guarantees that the banking system will make a return. The larger the profit rate, the more profit the bank will receive, and the more stable the bank will be.

Figure 2 shows the change in the amount of banks in the banking system and the associated systemic risk, and the parameter is set to $r_a = 0.004$, $r_b = 0.008$, $\chi = 0.3$, $\beta = 0.25$, $\overline{A} = 1000$, $\overline{\omega} = 500$, $\sigma_a = 0.15$, $\sigma_\omega = 0.25$, $C = 0.03$, $p = 0.05$, $\tau = 0.008$, $\theta = 0.02$, $\bar{r} = 0.008$. The trend in the amount of surviving banks over time in Figure 2 shows that in the early stages of the simulation, due to the interdependence of interbank assets and liabilities, the systemic risk was higher, leading to the rapid failure of some banks and a steep decrease in the amount of surviving banks. At the same time, systemic risk simulations decline rapidly at the initial stage, due to the fact that the initial situation of banks in the banking system is different, and the investment strategies developed are also different, and the systemic risk is greatest at this point. As it evolved, the banking system began to adapt and absorb some systemic risks, and gradually stabilized. In the realm of real finance, investment behavior itself contains risks. For banks with small and illiquid net worth, it is easy to fall into bankruptcy due to poor investment strategies. Failed banks may be incapable of repaying interbank loans in full, and overlapping bank portfolios can cause the value of other banks' assets to shrink, triggering a domino effect of bank failures. For some banks in good initial condition, if the investment has a poor strategy and yield is low, it may eventually lead to a continuous decline in profits and bankruptcy.

Figure 3 shows the evolution process of various parameters of a bank based on the risk contagion model constructed in this article. The parameter settings are consistent with Figure 2. As shown in the figure, the parameters of the bank's ownership equity, investment, dividends and liquidity change with the changes of internal and external factors in the banking system, and are in a state of dynamic change. When the ownership equity of the bank is greater than 0,

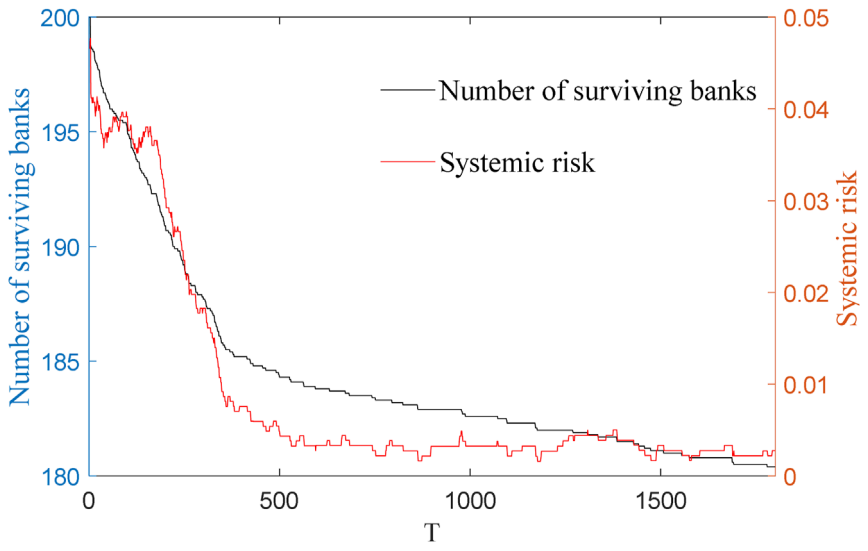


FIGURE 2
Evolution of the number of banks and systemic risks in the banking system.

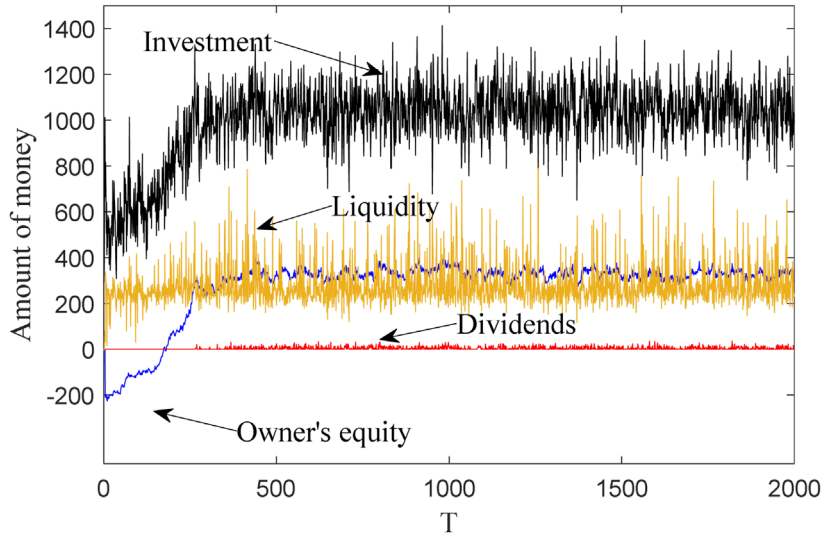


FIGURE 3
The evolution of the bank's internal parameters.

the bank's investment and dividends begin to increase, which is consistent with the actual operation of the bank.

This paper will further validate the validity of the model by analyzing the impact of parameters such as savings interest rate, savings volatility, reserve ratio, interbank lending rate, investment yield and volatility on systemic risk.

3.1 The impact of savings-related factors on systemic risk

Savings are essential for banks, not only as they are the main source of funding for banks, but also provide liquidity support to banks and are the basis for banks to make loans and other investment activities. Banks increase the value of their funds by paying interest on their savings to attract deposits, which are then used to make loans. The level of savings interest rate directly affects the bank's cost of funds, which in turn affects the bank's profitability and risk management.

Figure 4a illustrates the development of banks' systemic risk under different savings interest rate conditions. The parameter settings are consistent with Figure 2 except for r_a. As the figure shows, the savings interest rate has a significant impact on the systemic risk of banks; the lower the savings interest rate, the more stable the banking system, and the systemic risk of banks increases as the savings interest rate rises. When the savings interest rate is

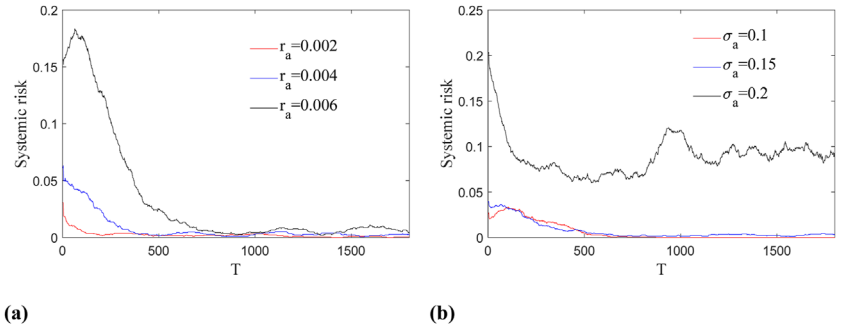


FIGURE 4
Evolution of systemic risk under different savings-related parameters: **(a)** Evolution of systemic risk under different savings rates; **(b)** Evolution of systemic risk under different savings volatility.

higher, the bank needs to pay more interest, which will raise the bank's funding cost, decrease the bank's profit margin, and thus increase the bank's operational risk. In addition, high interest rates can lead to the loss of deposits, further exacerbating banks' liquidity risks and ultimately leading to the failure of some banks that are not in good shape.

Figure 4b illustrates the evolution of systemic risk under different savings volatility conditions. The parameter settings are consistent with Figure 2 except for σ_a. It can be observed that when savings volatility is 0.1 and 0.15, the banking system eventually reaches a steady condition. When the savings volatility is 0.2, systemic risk is always at a high level. In reality, an increase in savings volatility means that banks face less deposit stability, which can lead to maturity matching issues for banks' balance sheets. When market interest rates fluctuate greatly, depositors are more inclined to transfer funds to other investment products with higher returns, thus affecting the structure of banks' funding sources. This flow of funds not only affects the liquidity of the bank, but can also result in a mismatch between the maturity of the assets and liabilities, increasing the bank's liquidity risk and interest rate risk.

3.2 The impact of the reserve requirement ratio on systemic risk

The reserve requirement ratio is the lowest ratio of reserves that the central bank demands to be held by commercial banks and is main tools of monetary policy. In the early days of the statutory reserve system, its main function was risk management. By setting a statutory reserve ratio, the central bank mandates that commercial banks retain a fixed percentage of cash reserves to meet the withdrawal needs of depositors. This system has effectively reduced the incidence of bank runs and preserved the banking system's stability [31]. Typically, a higher reserve requirement ratio (RRR) can enhance a bank's resilience to risk, as a higher reserve deposit means that the bank has more funds to deal with possible depositor withdrawals and liquidity pressures, reducing the likelihood of a run on the bank.

Figure 5 shows the evolution of systemic risk in various RRR states. The parameter settings are consistent with Figure 2 except for β. The simulation result shows that either too high or too low a reserve ratio increases the systemic risk of banks. It is clear from the figure that before 500 time-steps, the systemic risk increases as the RRR increases. On the one hand, higher reserve ratios require banks to keep more funds as reserves to cover depositor withdrawals and short-term liquidity needs. This enhances banks' ability to respond to runs and liquidity crises and reduces liquidity risk. On the other hand, excessively high reserve ratios reduce the funds available to banks for investment and lending, lowering investment income and interest income from loans, which in turn compresses profit margins and weakens banks' profitability and competitiveness. Under such circumstances, banks may adopt more risky investment strategies to make up for lost profits, increasing their own risk exposure, which may be transmitted to other banks through the financial network, exacerbating systemic risk. After 500 steps, the systemic risk at $\beta = 0.15$ exceeds the systemic risk at $\beta = 0.25$, because when the reserve ratio is too small, the bank's funds for investment increase, and the investment is risky, once the bank suffers a large loss, it is easy to go bankrupt. Therefore, policymakers need to consider when adjusting RRRs, and the need for further research on the specific effects of RRRs on systemic risk in different economic environments, so as to find a balance between enhancing bank stability and promoting economic growth.

3.3 The impact of interbank lending rates on systemic risk

The lending rate is a major factor influencing the liquidity and stability of banks. In financial markets, banks manage their short-term funding needs through lending markets to ensure they can meet customer withdrawal needs and other short-term obligations. Figure 6 shows the evolution of banks' systemic risk at different lending rates. The parameter settings are consistent with Figure 2 except for r_b. From the figure, it is obvious that with the increase of lending rates, bank systemic risk seems to have little change, that is, lending rate hardly affects the stable operation of the banking system. In June 2013, there was a turning point in China's financial history, known as the "money shortage". During this period, there was an extremely tight situation in the market, and all financial entities followed the principle of "cash is king" and refused to lend funds. On June 20, the Shanghai Interbank Offered Rate (SHIBOR)

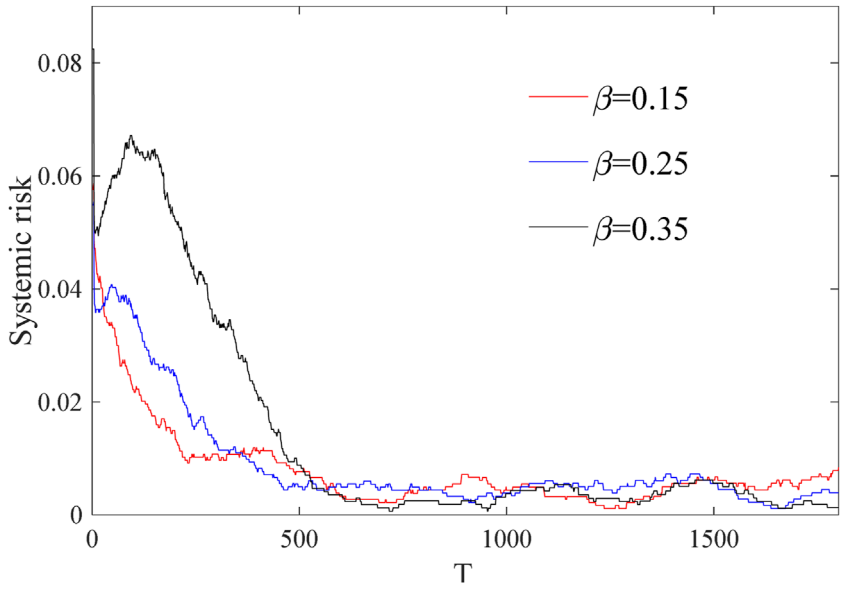


FIGURE 5
Evolution of systemic risk under different reserve ratios.

reached 13.44% [32], and the interbank market pledged repo rate (R001) hit a record high of 30%, but in the end, there was no bank bankruptcy, consistent with the simulation results. One reason was that the impact of short-term interbank lending rate shocks could be absorbed through investment returns, and the other was that the central bank injected 360 billion yuan in liquidity to prevent risk contagion. Similar cases include the 2008 financial crisis, when the United States 3-month LIBOR reached 4.82% but no large-scale bank failures occurred, and the 2011 European debt crisis, when the Euribor rose to 5.2% but core economies did not collapse. This indicates that, under an effective policy environment, the direct impact of interbank lending rates on systemic risk can be ignored. However, their sustained high levels may signal underlying pressures that require policy responses. The impact of investment-related parameters on systemic risk.

In the real financial market, the rate of return on investment is an important indicator to measure the profitability of a bank's investment activities. It reflects the ratio between the income that a bank earns through its investment activities and the cost of its investment. A higher yield usually means that the bank's investment activities are more successful, which can lead to a more stable source of income for the bank, thereby enhancing the bank's profitability and stability. Conversely, a lower yield means that the bank's investment activities are subject to higher risk, which may affect the stability of the bank.

Figure 7a shows the evolution of banks' systemic risk under various average rates of return. The parameter settings are consistent with Figure 2 except for τ. It is evident that the higher the average rate of return, the lower the banks' systemic risk. Over the entire period, the systemic risk at $\tau = 0.01$ decreases the fastest and is consistently lower than the systemic risk values at $\tau = 0.008$ and $\tau = 0.006$, and the banking system reaches stability at 500 steps, while the systematic risk is always maximum at $\tau = 0.006$. This is due to the fact that when the return on investment is large, a large interest rate differential can be generated between investment profits and savings interest, which can improve the bank's profitability and liquidity, thereby improving the bank's ability to resist risks and reduce systemic risks.

Figure 7b illustrates banks' systemic risk under different ROI fluctuations. The parameter settings are consistent with Figure 2 except for θ. It is clear that in the initial phase, the systemic risk measures were high for all curves, but over time, these indicators showed a downward trend. Throughout the time period, the systemic risk at ROI of 0.02 and 0.04 decreased rapidly, and the banking system was relatively stable, while the banking system was always exposed to higher systemic risk when the ROI was 0.06. This result shows that lower volatility of return on investment can help reduce the banking system's systemic risk, because when the volatility of investment is low, the investment income of banks is more stable, and the possibility of excessive losses in investment activities is less, which means that banks can effectively control risks and the banking system is more stable. When the investment volatility is high, the bank is more likely to suffer investment losses, and if the bank is not in a good state at this time, it can easily lead to the bank's bankruptcy.

3.4 The impact of bank size on systemic risk

In complex financial networks, bank size is not only a measure of the size of a single institution, but also a key hub for the propagation of systemic risk. In this paper, banks are categorized into different types according to their size, where those with an initial ownership interest of less than 100 are small-sized banks, those with an initial ownership interest between 100 and 500 are

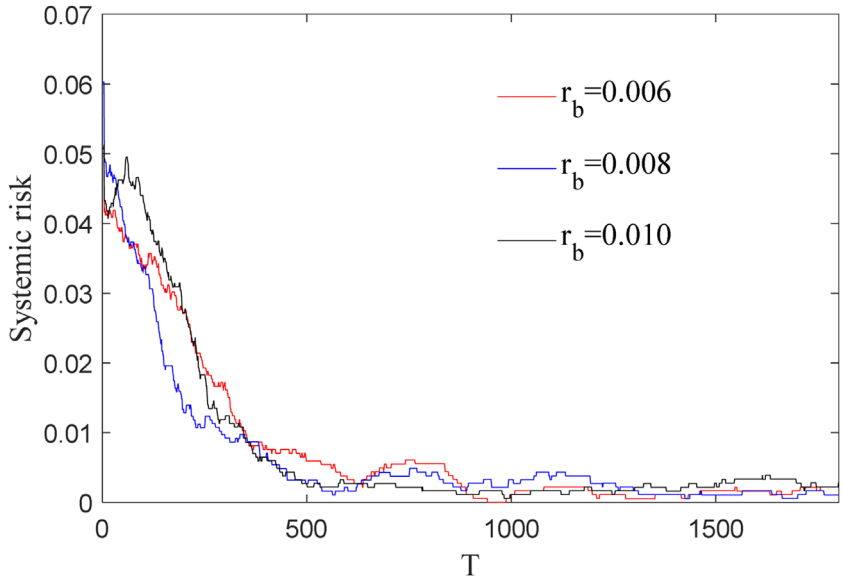


FIGURE 6
Evolution of systemic risk under different lending rates.

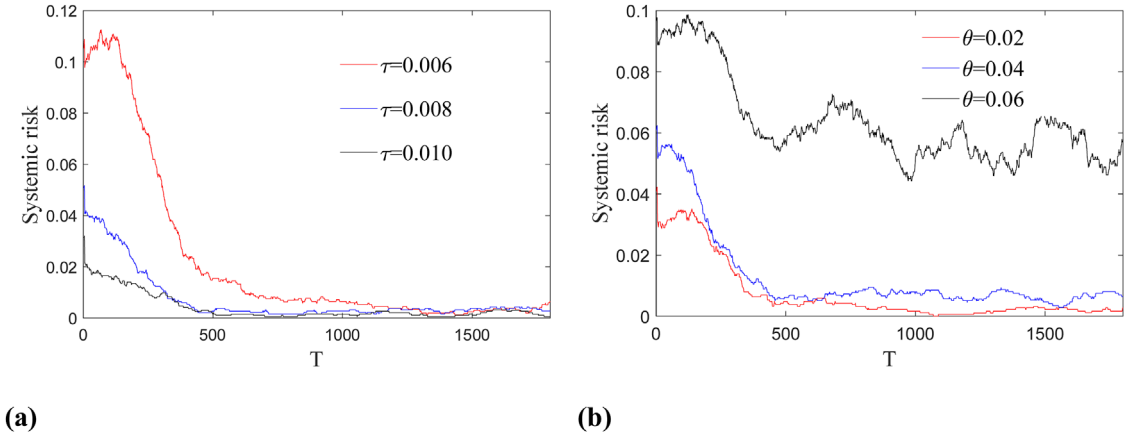


FIGURE 7
Evolution of systemic risk under different investment-related parameters: **(a)** Evolution of systemic risk under different average returns; **(b)** Evolution of systemic risk under different ROI fluctuations.

TABLE 1 Number of different types of banks and proportion of failed banks.

Type of bank	Number of banks	Percentage of failed banks
Small-scale bank	74	20.65%
Medium-scale bank	91	1.69%
Large-scale bank	35	0.00%

medium-sized banks, and those with an initial ownership interest greater than 500 are large-sized banks. The number of different types of banks and the percentage of failed banks are given in Table 1.

From the table it can be seen that small scale banks have the highest number of failures followed by medium scale banks and large-scale banks have almost no failures. This phenomenon is consistent with the "too big to fail" hypothesis. In fact, large banks are highly systemically important due to their size and influence [33], and are unable to cope with the negative knock-on effects of failure, so they will tide over the difficulties with government support. 2008 financial crisis, Lehman Brothers, due to over-investment in financial derivatives related to subprime bonds in the real estate market bubble, has become a major player in the financial crisis. Subprime debt-related financial derivatives, and suffered huge losses when the real estate market collapsed, and eventually filed for bankruptcy protection on September 15, 2008. Therefore, despite the low rate of large-scale bank failures, cross-asset investments and

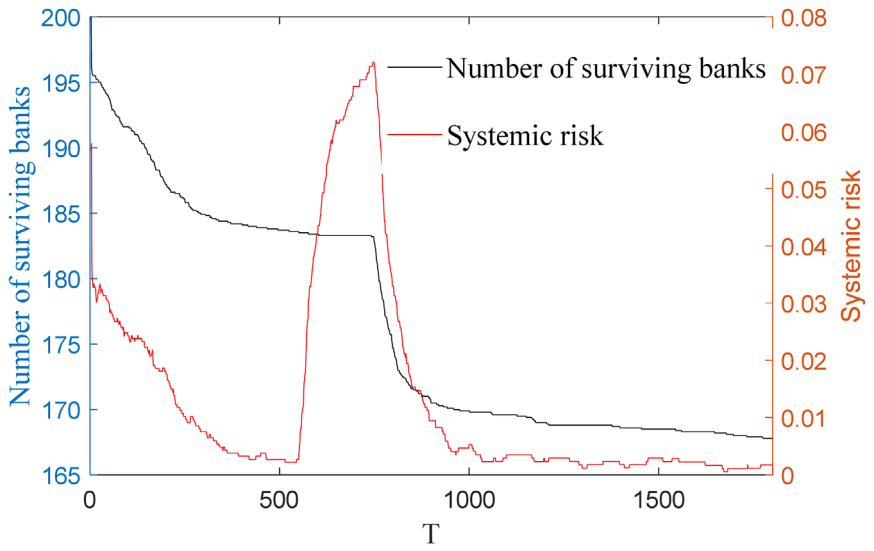


FIGURE 8
Evolutionary outcomes of bank networks when asset prices fall.

TABLE 2 The impact of key parameters on systemic risk.

Parameter name	Parameter values	Systemic average	Risk stability[a]
Savings rates (r_a)	0.002	0.0043	1
	0.004	0.0098	1
	0.006	0.0536	0
Savings volatility (σ_a)	0.1	0.0069	1
	0.15	0.0090	1
	0.2	0.0830	0
Reserve ratios (β)	0.15	0.0107	1
	0.25	0.0083	1
	0.35	0.0130	1
Average returns (τ)	0.006	0.0293	0
	0.008	0.0096	1
	0.01	0.0044	1
ROI fluctuations (θ)	0.02	0.0103	1
	0.04	0.0117	1
	0.06	0.0661	0

[a]The system eventually reaches a steady state is denoted by 1, otherwise it is denoted by 0.

subprime exposures need to be monitored. On the other hand, the likelihood of a financial crisis depends on the liquidity position of banks in the system [34], so small and medium-sized banks need improved liquidity support to reduce the likelihood of outright bankruptcy.

3.5 Impact of falling asset prices on systemic risk

In order to assess the robustness of the banking system to extreme shocks, this study designs an asset price crash scenario with reference to historical financial crisis characteristics. Figure 8 presents the evolutionary results of the banking network suffering a 40% fall in asset prices at step 750, with the same parameter settings as in Figure 2. From the figure, it can be seen that the number of insolvent banks increases rapidly after suffering an asset price fall, and due to the high degree of correlation in the banking system, the asset price fall transmits risk among banks through multiple channels, which leads to a domino effect of a large number of banks failing one after another.

The results in Figure 8 show that systemic risk, after a short-term sharp rise, gradually decreases to a steady state, the number of bank failures gradually decreases, and the banking system stabilizes, indicating that the banking network is able to gradually absorb systemic risk.

On the basis of the foregoing analysis, it may be concluded that the banking system's stability is related to a variety of factors. Table 2 demonstrates the impact of several key parameters on systemic risk, and here we analyze it in the context of the specific economic situation. When the economy is growing steadily, the banking system is generally exposed to low systemic risk. In this environment, investor yields are usually higher and less volatile, helping banks maintain stable profitability and thus enhancing their stability. In times of increased economic uncertainty or economic downturn, banks may be exposed to higher systemic risks, especially when the reserve requirement ratio is high, which may limit banks' lending capacity and affect their profitability and liquidity. At this

juncture, an appropriate reduction in the reserve requirement ratio and the savings interest rate is conducive to replenishing banks' liquidity and improving their ability to withstand risks. Therefore, banks and policymakers need to pay close attention to changes in these factors and take appropriate risk control measures to preserve the banking system's stability.

4 Discussion

The financial system is a complicated network structure consisting of diversified financial entities and their interactions, which include not only the direct relationship between interbank loans and cross-ownership of subordinated debt, but also the indirect coupling effect caused by the convergence of asset allocation. It is of great significance for regulators and bank policymakers to study the impact of bank savings interest rates and their fluctuations, interbank lending rates, reserve requirement ratios, and investment yields and fluctuations on bank stability. Previous studies have mostly focused on the coupling of interbank lending and investment, and few have taken into account cross-subordinated debt holdings. In addition, it is difficult for the existing model to truly reflect the dynamic development of the banking system because some assumptions in the existing model are not realistic, such as the same interest rate on subordinated bonds issued by each bank, and the risk appetite of the bank is not taken into account. To better reflect the development of the banking system, this paper establishes a multichannel risk contagion model from the risk appetite perspective, which comprehensively considers interbank loans, investment coupling, and the holding of subprime bonds with different interest rates according to different risk appetite levels. Based on multi-dimensional simulation and theoretical deduction, this paper draws the following conclusions.

(1) The savings interest rate can directly affect the bank's cost of funds, which in turn affects the profitability and stability of the bank. The higher the savings interest rate of the bank, the higher the fluctuations of savings and the higher the systemic risk. Appropriately lowering the interest rate on savings and reducing the fluctuation of savings is conducive to improving the stability of banks.
(2) A reserve requirement ratio that is too high or too low will raise the systemic risk of banks. If the reserve requirement ratio is too high, it will reduce the interest rate spread between investment and savings, and concurrently limit the bank's lending ability, affect the bank's profitability, and increase systemic risk. The smaller the interest rate on deposits, the more money the bank uses for investment, and the higher the bank's risk.
(3) The interbank lending rate is not a significant factor in the systemic risk of banks.
(4) The higher the average return on investment, the more stable the bank will be. If the bank's investment return is too low, the bank will not be able to make sufficient profits, which will have the effect of reducing liquidity and weakening the ability of the bank to resist risks. The less volatile the investments, the lower the bank's systemic risk. A sound investment strategy is also essential to improve the stability of the bank.

Based on the above conclusions, we propose the following policy recommendations: For regulatory authorities, it is necessary to improve the monitoring indicator system for banking system stability, fully considering the impact of factors such as savings interest rates, reserve ratios, investment returns, and volatility on systemic risk, closely monitoring trends in these indicators, and promptly identifying potential risks. When formulating and adjusting reserve requirement ratio policies, the dual impact on systemic risks in the banking system should be carefully balanced. Differentiated reserve requirement ratio policies can be implemented based on the scale and risk characteristics of different types of banks. Additionally, an emergency mechanism for the interbank lending market should be established and improved, with enhanced monitoring of the flow of interbank lending funds. For banks themselves, savings products should be priced reasonably, with savings interest rates determined scientifically and reasonably based on their cost structures, risk preferences, and market competition conditions. Banks should also optimize their asset allocation strategies, focusing on diversification and balancing risk and return, while reasonably controlling the volatility of investment assets. Furthermore, banks should strengthen liquidity risk management, reasonably arrange funding sources and usage, ensure sufficient liquidity to address funding shortages, and enhance communication and coordination with the central bank and other financial institutions.

This study constructs a dynamic evolution model framework, which aims to quantitatively measure and analyze the mechanism of bank systemic risk in the context of macroeconomic cyclical fluctuations, and provides methodological support for regulators to identify risk contagion paths and formulate forward-looking prevention and control strategies. Future research directions can focus on the following aspects: first, calibrating model parameters based on real transaction data to improve the empirical adaptability of risk prediction; Second, more heterogeneous bank behavior assumptions (such as differentiated liquidity management strategies) are introduced to enhance the explanatory power of the model to the complex financial ecology. This will further improve the dynamic monitoring framework for systemic risks and provide a science-based foundation for the precise implementation of macroprudential policies.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/knochen336/Risk-Contagion-of-Banks-Holding-Subordinated-Debt/.

Author contributions

SC: Data curation, Writing – original draft, Methodology, Software. SJ: Project administration, Writing – review and editing, Conceptualization, Funding acquisition, Supervision. YZ: Investigation, Formal Analysis, Data curation, Writing – review and editing. LM: Writing – review and editing, Visualization, Investigation, Validation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is funded by Ministry of Education, Humanities and social science projects (21YJC790054) and National Natural Science Foundation of PR China (72101121).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Poledna S, Martínez-Jaramillo S, Caccioli F, Thurner S. Quantification of systemic risk from overlap portfolios in the financial system. *J Financial Stab* (2021) 52:100808. doi:10.1016/j.jfs.2020.100808

2. Jackson MO, Pernoud A. Systemic risk in financial networks: a survey. *Annu Rev Econ* (2021) 13(1):171–202. doi:10.1146/annurev-economics-083120-111540

3. Craig B, Ma Y. Intermediation in the interbank lending market. *J Financial Econ* (2022) 145(2):179–207. doi:10.1016/j.jfineco.2021.11.003

4. Qi M, Shi D, Feng S, Wang P, Nnenna AB. Assessing the interconnectedness and systemic risk contagion in the Chinese banking network. *Int J Emerging Markets* (2022) 17(3):889–913. doi:10.1108/ijoem-08-2021-1331

5. Liu X. Unraveling systemic risk transmission: an empirical exploration of network dynamics and market liquidity in the financial sector. *J Knowledge Economy* (2024) 16:6629–64. doi:10.1007/s13132-024-01861-9

6. Macchiati V, Brandi G, Di Matteo T, Paolotti D, Caldarelli G, Cimini G. Systemic liquidity contagion in the European interbank market. *J Econ Interaction Coord* (2022) 17(2):443–74. doi:10.1007/s11403-021-00338-1

7. Yanquen E, Livan G, Montañez-Enriquez R, Martinez-Jaramillo S. Measuring systemic risk for bank credit networks: a multilayer approach. *Latin Am J Cent Banking* (2022) 3(2):100049. doi:10.1016/j.latcb.2022.100049

8. Capponi A, Sun X, Yao DD. A dynamic network model of interbank Lending—Systemic risk and liquidity provisioning. *Math Operations Res* (2020) 45(3):1127–52. doi:10.1287/moor.2019.1025

9. Gao Q, Lv D, Jin X. Systemic risk of multi-layer financial network system under macroeconomic fluctuations. *Front Phys* (2022) 10:943520. doi:10.3389/fphy.2022.943520

10. Ding Z, Yan H, Chen Y, Jiang C. Risk contagion in interbank lending networks: a multi-agent-based modeling and simulation perspective. *Expert Syst Appl* (2024) 256:124847. doi:10.1016/j.eswa.2024.124847

11. Fan H, Sheng W. Research on dynamic evolution of systemic risk in Chinese banking system. *Front Phys* (2020) 8:199. doi:10.3389/fphy.2020.00199

12. Zou J, Fu X, Yang J, Gong C. Measuring bank systemic risk in China: a network model analysis. *Systems* (2022) 10(1):14. doi:10.3390/systems10010014

13. Jin S, Song L, Shu L, Gao Q, Chen Y. Systemic risk in Chinese interbank lending networks: insights from short-term and long-term lending data. *Empirical Econ* (2024) 67:2539–64. doi:10.1007/s00181-024-02617-9

14. Krause A, Giansante S. Interbank lending and the spread of bank failures: a network model of systemic risk. *J Econ Behav and Organ* (2012) 83(3):583–608. doi:10.1016/j.jebo.2012.05.015

15. Li S, Liu M, Wang L, Yang K. Bank multiplex networks and systemic risk. *Physica A: Stat Mech Its Appl* (2019) 533:122039. doi:10.1016/j.physa.2019.122039

16. Wang H, Li S. Optimization of systemic risk: reallocation of assets based on bank networks. *J Risk* (2021). doi:10.21314/JOR.2020.449

17. Huang Y, Liu T. Diversification and systemic risk of networks holding common assets. *Comput Econ* (2023) 61(1):341–88. doi:10.1007/s10614-021-10211-9

18. Ma J, He J, Liu X, Wang C. Diversification and systemic risk in the banking system. *Chaos, Solitons and Fractals* (2019) 123:413–21. doi:10.1016/j.chaos.2019.03.040

19. Zou L, Xie L, Yang Y. A double-layer network and the contagion mechanism of China's financial systemic risk. *J Artif Societies Soc Simulation* (2019) 22(4). doi:10.18564/jasss.4122

20. Gao Q, Fan H. Systemic risk caused by the overlapping portfolios of banks under a bilateral network. *Front Phys* (2021) 9:638991. doi:10.3389/fphy.2021.638991

21. Caccioli F, Barucca P, Kobayashi T. Network models of financial systemic risk: a review. *J Comput Soc Sci* (2018) 1:81–114. doi:10.1007/s42001-017-0008-3

22. Nguyen T. The disciplinary effect of subordinated debt on bank risk taking. *J empirical Finance* (2013) 23:117–41. doi:10.1016/j.jempfin.2013.05.005

23. Capponi A, Weber M. Systemic portfolio diversification. *Operations Res* (2024) 72(1):110–31. doi:10.1287/opre.2022.0290

24. Shen P, Li Z. Financial contagion in inter-bank networks with overlapping portfolios. *J Econ Interaction Coord* (2020) 15(4):845–65. doi:10.1007/s11403-019-00274-1

25. Weber S, Weske K. The joint impact of bankruptcy costs, fire sales and cross-holdings on systemic risk in financial networks. *Probab Uncertainty Quantitative Risk* (2017) 2:9–38. doi:10.1186/s41546-017-0020-9

26. Wang C, Liu X, He J. Does diversification promote systemic risk? *North Am J Econ Finance* (2022) 61:101680. doi:10.1016/j.najef.2022.101680

27. Jiang S, Fan H. Systemic risk in the interbank market with overlapping portfolios and cross-ownership of the subordinated debts. *Physica A: Stat Mech its Appl* (2021) 562:125355. doi:10.1016/j.physa.2020.125355

28. Yan C, Ding Y, Liu W, Liu X, Liu J. Multilayer interbank networks and systemic risk propagation: evidence from China. *Physica A: Stat Mech its Appl* (2023) 628:129144. doi:10.1016/j.physa.2023.129144

29. Fan H, Pan H. The effect of shadow banking on the systemic risk in a dynamic complex interbank network system. *Complexity* (2020) 2020(1):1–10. doi:10.1155/2020/3951892

30. Caccioli F, Shrestha M, Moore C, Farmer JD. Stability analysis of financial contagion due to overlapping portfolios. *J Banking and Finance* (2014) 46:233–45. doi:10.1016/j.jbankfin.2014.05.021

31. Pan H, Fan H. The stability of banking system with shadow banking on different interbank network structures. *Discrete Dyn Nat Soc* (2021) 2021(1):1–15. doi:10.1155/2021/6650327

32. Chen B, Li L, Peng F, Anwar S. Risk contagion in the banking network: new evidence from China. *North Am J Econ Finance* (2020) 54:101276. doi:10.1016/j.najef.2020.101276

33. Zhang X, Zhang X, Lee CC, Zhao Y. Measurement and prediction of systemic risk in China's banking industry. *Res Int Business Finance* (2023) 64:101874. doi:10.1016/j.ribaf.2022.101874

34. Chen H, Wang T, Yao DD. Financial network and systemic Risk—A dynamic model. *Prod Operations Manag* (2021) 30(8):2441–66. doi:10.1111/poms.13384

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 20 August 2025
DOI 10.3389/fphy.2025.1625934

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Asaduddin Abdullah,
IPB University, Indonesia
Yang Li,
Tsinghua University, China

*CORRESPONDENCE
Qiaoming Hou,
houqiaoming@sut.edu.cn

RECEIVED 09 May 2025
ACCEPTED 14 July 2025
PUBLISHED 20 August 2025

CITATION
Yang L and Hou Q (2025) Blockchain-driven logistics provider guaranteed financing strategy considering risk aversion.
Front. Phys. 13:1625934.
doi: 10.3389/fphy.2025.1625934

COPYRIGHT
© 2025 Yang and Hou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Blockchain-driven logistics provider guaranteed financing strategy considering risk aversion

Lei Yang and Qiaoming Hou*

School of Management, Shenyang University of Technology, Shenyang, China

Introduction: In e-commerce supply chains, suppliers generally face funding constraints, and the guaranteed financing provided by logistics providers can effectively improve the capital liquidity of suppliers. Blockchain helps alleviate the information silos in the financing process, strengthen cooperation between enterprises, and enhance the quality of financing. Cooperative and non-cooperative decisions have significant impacts on the formulation of financing strategies, and contract design helps coordinate the interest conflicts between enterprises and optimize financing strategies.

Methods: Addressing the financial constraints for suppliers, this study employs game theory analysis to construct a guarantee financing model under blockchain technology where logistics providers offer guarantees to suppliers. It considers the risk-averse of managers and first analyzes the financing equilibrium strategies before and after blockchain in a non-cooperative game setting. Then, under a cooperative game framework, it explores the financing equilibrium strategies in three modes: cooperation between suppliers and logistics providers, cooperation between logistics providers and e-commerce platforms, and cooperation among all three parties. It also compares and analyzes the influence of blockchain on cooperative and non-cooperative game strategies in terms of financing decisions. To coordinate the interest conflicts among various risk-averse members, the study introduces revenue-sharing and cost-sharing to formulate coordination strategies.

Results and discussion: Through comparative analysis, it is found that information verification costs can impair the financing gains of suppliers but help ensure the financing gains of logistics providers. Supplier defaults are detrimental to the development of supply chain financing (SCF). The financing gains of suppliers and logistics providers will initially decrease and then increase with the adoption of blockchain, which can better optimize financing strategies once they exceed a certain threshold. A higher risk aversion can harm the financing gains of suppliers and logistics providers but benefit the gains of ecommerce platforms and financial institutions. Compared to non-cooperative modes, cooperative models are more effective in enhancing the financing efficiency, and under cooperative modes, blockchain is more conducive to promoting the improvement of corporate financing gains. The Shapley value can effectively allocate the financing gains among supply chain members, and risk-sharing contracts can increase the financing gains of suppliers and logistics

providers. This research aims to provide theoretical references for the financing decisions of supply chain enterprises.

KEYWORDS

blockchain, risk aversion, supply chain financing, financing equilibrium, strategy optimization

1 Introduction

In the e-commerce supply chain, the guarantee financing provided by logistics providers can primely relieve the financial trouble of suppliers. Most suppliers are small and medium-sized enterprises (SMEs), and their funding constraints are becoming increasingly prominent. Logistics providers, offering transportation services, have a thorough understanding of logistics, commerce flows, and other aspects based on their transactions with suppliers. They are more willing to provide guarantees to help suppliers obtain working capital in advance, thereby effectively easing their financial constraints. Under the guarantee of logistics providers, suppliers do not need to wait for the sale of goods to recover funds; instead, they can obtain short-term funds to pay for raw material procurement, production operations, and other expenses, significantly shortening the capital turnover cycle [1]. Logistics providers, leveraging their real-time supervision capabilities of goods and credit endorsements, reduce the lending risks for financial institutions, enabling suppliers to access financing at lower costs while assuring the stability of the supply chain's upstream and downstream segments [2]. Through the guarantee of logistics providers, banks would provide loans to suppliers. This model not only optimizes the cash flow management of suppliers but also activates assets and enhances overall operational efficiency. Through cooperation among various members, it strengthens communication and interest coordination between enterprises, improves the efficiency of the use of funds, and optimizes financing strategies [3].

In the financing decisions, enterprises are increasingly focusing on risk avoidance. Although the financing model provided by logistics providers alleviates the financial pressure on suppliers, its potential risks still require high vigilance. The value of products is easily affected by market price fluctuations, product obsolescence, or damage, which may lead to delayed sales and the inability to timely recover funds, the supplier lacks sufficient funds to repay the loan [4]. Enterprises have a strong tendency towards risk avoidance, and managers should strengthen risk control through multiple means to address these risks. By introducing third-party evaluation agencies to dynamically monitor enterprise value, combining the Internet of Things (IoT) to achieve real-time tracking and data transparency of commerce and capital flows, and establishing dynamic adjustment mechanisms to set flexible financing limits based on market supply and demand and product characteristics to effectively avoid the risks faced by enterprises [5]. Supply chain enterprises are collaborating to build mechanisms for shared benefits and shared risks and using big data analysis to predict industry risks and develop contingency strategies in advance. These measures aim to balance financing efficiency with controllable risks, assuring the steady running of SCF ecosystem [6].

Blockchain technology can optimize financing strategies and coordinate interest conflicts among enterprises. Blockchain technology, through the construction of a trustworthy distributed ledger and smart contract mechanisms, can deeply optimize financing strategies and strengthen collaborative cooperation among supply chain enterprises. In the financing process, blockchain records the status and flow of funds in real time, ensuring information transparency and traceability, reducing trust costs for all parties, and providing precise data support for dynamic adjustments of guaranteed rates, financing limits, and other strategies, improve the availability of information [7]. Smart contracts can automatically execute pre-set rules for shared benefits and shared costs, automatically allocating the profits of logistics providers and suppliers according to predetermined ratios, and triggering risk compensation mechanisms based on market price fluctuations to balance interest conflicts among all parties, it helps to better alleviate the financial constraints of enterprises [8]. This technology-empowered contract coordination strategy not only avoids information games and execution lags in traditional negotiations but also promotes enterprises to actively maintain asset security and supply chain stability through incentive-compatible designs, improve financing and operational efficiency. A blockchain-driven financing ecosystem integrates dispersed participants into a community of shared interests, enhancing financing efficiency while building a new paradigm of supply chain finance characterized by controllable risks and win-win cooperation [9].

Based on this, we construct non-cooperative and cooperative game models, consider the risk-averse behavior of managers, and introduce revenue-sharing and risk-sharing as coordination mechanisms to explore the equilibrium and coordination strategies of guarantee financing provided by logistics providers under blockchain. The main issues addressed are: (1) What are the financing equilibrium strategies before and after the adoption of blockchain? (2) How to coordinate the interest conflicts among various supply chain enterprises? The contributions are mainly reflected as follows: Firstly, applying blockchain to the guarantee financing model provided by logistics providers, it compares and analyzes non-cooperative and cooperative games for the first time, exploring financing equilibrium strategies under different scenarios. Previous studies mostly considered applying blockchain technology to traditional supply chain financing models and were largely based on case studies or qualitative analyses. This study applies blockchain to a specific financing area and uses game models to explore optimal financing equilibrium strategies. Secondly, in the financing decision-making process, it considers the risk-averse behavior of enterprises and analyzes the best finance strategies from the managerial behavior. Previous scholars, when studying financing strategies, would consider the financing risk issues faced by enterprises but rarely approached it from the perspective of managerial behavior, considering the risk-averse behavioral preferences of enterprises. This study incorporates the risk-averse of managers into the financing model to analyze the impact of

risk-averse behavior. Finally, it introduces revenue-sharing and risk-sharing to harmonize the conflicts of risk-averse supply chain enterprises under blockchain. Previous research mostly introduced contract coordination strategies in the study of supply chain management, and coordination strategies also have big impacts on the finance decisions. This study introduces revenue-sharing and risk-sharing in the analysis of financing strategies to coordinate interest conflicts among enterprises.

The subsequent structure arrangement is as follows: The second part elaborates on the current research status of existing literature. The third part describes the background of the research problem and proposes relevant hypotheses. The fourth part constructs a non-cooperative game before and after blockchain. The fifth part constructs a cooperative game model after blockchain. The sixth part analyzes the coordination strategies after cooperation. The seventh part investigates the impact of various factors through numerical analysis. The eighth part is the summary and outlook.

2 Literature review

2.1 Guarantee financing in SCF

Logistics provider guaranteed financing, as a crucial component of SCF, plays a great role in addressing capital constraints.

In terms of influencing factors, the guaranteed financing by logistics providers can effectively alleviate the funding constraints of enterprises. The scale of 3PL significantly affects banks' willingness to grant credit and the financing decision mechanism. Large-scale 3PLs can enhance the willingness to grant credit to SMEs, whereas small-scale 3PLs lead to higher perceived risks. Data results show that smaller scales induce larger N2, while larger scales trigger larger LPP [10]. Through interviews, the study explores how TPL companies with different strategic orientations leverage market practices to gain competitive advantages. The results indicate that scale-oriented enterprises focus on market expansion and view SCF as the best tool to enhance benefits; scope-oriented enterprises regard SCF as a core business and improve financial performance through SCF empowerment [11]. In a SCF system consisting of manufacturers, retailers, and 3PLs, demand is influenced by logistics service levels and retail prices. Stackelberg games are constructed separately for retailer or manufacturer funding constraints, revealing that when the retailer is constrained, the manufacturer's price keeps steadfast, while other participants' strategy and gain rise with higher logistics service sensitivity and decrease with improved cost efficiency. When the manufacturer is constrained, the strategy and gain of all parties similarly increase with higher logistics service susceptibility and decrease with improved cost efficiency. The study recommends that retailers and manufacturers should prioritize 3PL's service decisions under the 3PL financing model [12]. To investigate financing risks, an order financing risk assessment index system is established, combining letters and expert interviews to establish five dimensions: financing enterprises, downstream core enterprises, logistics enterprises, supply chain running, and environment. The indicators are perfected through questionnaires and statistical analysis and validated with an order finance project jointly undertaken by banks and logistics [13].

In terms of strategy selection, scholars have studied the equilibrium decisions of e-commerce platforms and 3PLs in providing guarantee financing and sales model choices for small and medium-sized suppliers. It was found that under different market risks and service costs, SMEs tend to opt for platform or 3PL guarantee financing, and guarantee financing influences the platform's choice between resale and market modes [1]. 3PLs provide credit to budget-constrained retailers through a financial integration model, which can strengthen the gain of 3PLs, suppliers, retailers, and the supply chain compared to supplier credit and the no-budget constraint model. When the marginal profit of 3PLs is higher than that of suppliers, this control model is superior to supplier credit, and this model can surpass the no-budget constraint model [14]. For capital-constrained retailers, a bankruptcy guarantee insurance contract is designed, and third-party insurance institutions are introduced to bear bankruptcy losses. Under two financing mechanisms—trade credit and bank credit—the optimal decisions for all parties are provided, and the impact of parameters on decisions and profits is analyzed, offering references for formulating scientific policies and advancing supply chain financing [15].

In terms of model optimization, 3PLs leverage their information advantages to provide guarantees and direct financing for capital-constrained SMEs, but the default risks brought by demand fluctuations affect the choice of financing models. Studies compare the preferences of 3PLs, manufacturers, and retailers under guarantee financing, direct 3PL financing, and trade credit, finding that 3PL financing strategies depend on the manufacturer's costs and the retailer's initial funds. Financing equilibrium may damage the interests of all parties due to inconsistent strategies, necessitating coordination mechanisms to enhance the overall profit of the whole chain and provide guidance for 3PLs in selecting financing models under resilience considerations [16]. The study compares three 3PL financing models: bank financing, EPF, and 4PF. It shows that EPF improves commission rates and enhances logistics services, while 4PF is determined by 4PL fees; the three models complement each other, and when capital opportunity costs are homogeneous, platforms and 4PLs tend to self-finance [17]. A dual-channel supply chain game model is constructed to compare the finance models of e-commerce platforms and 3PLs. Increasing ECP interest rates reduces profits but increases total revenues; 3PL financing incurs lower direct operating costs than ECPs engaging in price competition through distribution; ECP financing is superior to 3PL financing; 3PLs need to lower interest rates or differentiate freight charges to compete [18]. Using fsQCA and interview methods, the study explores the cost-reduction effects of SCF and traditional finance combinations. It shows that cost reduction effects are constrained by traditional financing methods, and enterprises using credit guarantee financing are more likely to reduce costs compared to guarantee financing. The study provides references for policymakers and practitioners to simplify applications [19]. Addressing the two financing options for manufacturers, bank credit financing (BCF) and supplier green investment (SGI), this study analyzes the optimal strategy choices and examines the effects of various supply chain coordination contracts. In terms of coordination mechanisms, a revenue-sharing cannot coordinate enterprises, but both cost-sharing and quantity discount can effectively coordinate them when appropriate cost-sharing ratios or quantity discount rates are set [20].

Scholars studying the guarantee financing strategies provided by logistics providers have mainly focused on analyzing influencing factors, strategy selection, and model optimization. There are multiple factors affecting the optimal financing strategy, with risk being the primary influencing factor. Contemporary research mostly considers the influence of external risks on finance strategies. Building on this, this study analyzes the company's own attitude towards risk avoidance and explores the influence of risk-averse on optimal finance decisions from the perspective of managerial behavior. In terms of strategy selection, scholars point out that different financing strategies should be chosen for different financing scenarios to optimize financing when formulating financing strategies. This study further introduces blockchain technology on this basis, utilizing its technological advantages to optimize financing strategies. It conducts a detailed analysis of the financing equilibrium before and after blockchain, exploring the optimal finance strategy. In terms of mode comparison, scholars compare the financing services provided by logistics providers with those offered by e-commerce platforms and other supply chain financing models, finding that logistics provider financing has better advantages. This study further analyzes the guarantee financing services provided by logistics providers, exploring the financing equilibrium strategies before and after the adoption of blockchain, and introduces a cooperative game model to investigate the influence of collaboration between logistics providers, e-commerce platforms, and other supply chain members on financing strategies. To optimize the financing model, this study introduces revenue-sharing and cost-sharing based on existing research to coordinate the interest conflicts among e-commerce supply chain members.

2.2 SCF under blockchain

Many advantages of blockchain have a good application in supply chain financing, which can effectively optimize the finance strategy.

In terms of operational mechanisms, scholars have proposed blockchain-based SCF models to address issues such as information asymmetry and cybersecurity risks in traditional platforms. The decentralized, transparent, and immutable nature of blockchain helps enhance transaction transparency and trust, protect the rights of SMEs, and optimize the efficiency of cash flow [21]. Blockchain can solve problems related to financing coverage, cost, efficiency, and risk control in SCF. Before implementation, it requires joint platform construction by multiple parties and organizational adjustments, with pilot feedback driving restructuring; to achieve scaled application, it needs to incentivize more participants and improve regulations. Based on the innovation theory, an implementation frame is put to provide guidance for system [22]. A blockchain logistics prototype system based on Hyperledger was designed and tested in real logistics enterprises, building a tracing platform based on logistics service flows to implement supplier responsibilities, improve LSP performance, and promote the development of Logistics 4.0, revealing the transformative benefits of blockchain in logistics services [23]. The supply chain payment process is cumbersome, and information is isolated, making it difficult to associate transportation with payments. A blockchain platform based on Ethereum smart contracts was constructed to achieve automatic authorization and information sharing. Suppliers can reduce the risk of consignment sales, buyers can track in real-time to ensure receipt of goods, and 3PLs can reduce paperwork and tracking costs, improving efficiency; the solution is secure and transparent [24]. E-commerce improves supply chain efficiency but leads to cash shortages for small retailers; existing logistics finance relies heavily on large 3PLs, with high thresholds. A blockchain-driven logistics finance platform is proposed, adopting a cross-layer architecture and smart contracts, combined with multi-agent systems, to achieve dynamic collateral management, and the feasibility is verified through case studies [25]. Integrating blockchain into transportation capacity supply chain financing, analyzing its impact on credit access, limits, and regulation. Through credit transmission decay, credit assessment, and closed-loop supervision, the mechanisms by which blockchain optimizes financing in three scenarios are revealed, indicating that it can enhance the participation of financial institutions and provide better credit to transportation enterprises [26].

In terms of strategy optimization, scholars have proposed a blockchain-based e-commerce logistics financing network physical traceability system, achieved logistics information visualization and shared through digital asset tokenization, promoting logistics financing operations using shared ledgers and smart contracts, and demonstrating system design and application, reducing upfront investment, enhancing flexibility, and supporting supply chain development [27]. Internet + policies have promoted the development of internet finance, with banks serving SMEs through SCF, combining logistics technology and edge computing to achieve real-time, secure data processing. A blockchain-based financing management system is studied, analyzing cash flow and risk control, and optimizing costs and efficiency through shared data and advanced processing, reducing risks [28]. Globalization and deepening production division have led to complex supply chain networks, making traditional models difficult to cope with. A blockchain-empowered SCF risk assessment and behavior prediction algorithm is proposed, constructing a dual-chain management system and architecture, with weighted risk control indicators based on homogeneity of variance testing, testing and comparing the performance of TPR, FPR, etc., for data from five companies, aiming to improve financing efficiency and prevent risks [29]. Small manufacturing enterprises are constrained by large financing needs and low trustworthiness. Traditional supply chain finance only benefits first-tier suppliers, while blockchain-driven innovations establish a trust transmission model through electronic payment vouchers, introducing dynamic weight incentives to expand multi-layer financing, alleviating finance trouble for SMEs, and the model's rationality is verified through numerical analysis [30]. Case studies explore how blockchain drives the reconstruction of financial platform models. The study finds that blockchain empowers information, funds, logistics, and value flows, with model reconstruction driven by internal (content, governance, transaction structure) and external (industry, region, technology) factors, and enhances value proposition, network expansion, and value creation and distribution [31]. Scholars have also explored the value creation process of blockchain-empowered SCF. Based on service-dominant logic and social exchange theory, the roles, motivations, resources, and practices of participants are analyzed, and through case studies of blockchain platforms, it is verified that it meets needs through key

resources and practices, creating value for all parties in the supply chain. It provides empirical evidence and implications [32].

In terms of financing model, a three-tier supply chain model for blockchain platform finance (BPF) is constructed, and through decision-making and parameter sensitivity analysis, the optimal strategy is identified. Compared to independent financing, it is found that when the retailer has insufficient capital and high costs, BPF is superior for manufacturers, distributors, and retailers. Risk sharing can enhance financing efficiency, providing theoretical basis for the selection of three-party financing models [33]. A three-tier supply chain game model is constructed to compare the effect of traditional and blockchain-visualized prepayment financing on risk mitigation and financial decision-making. The results show that blockchain improves visibility and refines financing decisions; consignment financing only achieves win-win for all parties when the second-tier supplier is severely capital-constrained; cross-layer direct financing achieves win-win for all three parties [34]. Based on the credibility of blockchain, SMEs can self-guarantee financing. A four-party model compares PUG with CUG, finding that CUG can bring Pareto improvement, but if the customer's opportunity cost is bigger than the platform's, the platform will be disadvantaged. The platform can adopt a high advance payment, low service fee strategy to reduce the risk of low-credit customers, and the advance payment ratio does not affect costs [35].

For blockchain-based supply chain financing, scholars have explored the impact mechanisms and optimization paths of blockchain on financing decisions, analyzing the factors influencing the financing profits of supply chain members. Comparative analyses of traditional financing decisions and blockchain-based financing decisions have found that blockchain helps to reduce financing risks. However, regarding the guarantee financing provided by logistics providers, fewer scholars have integrated the characteristics of the financing model, introducing the guarantee provided by logistics providers to investigate the financing equilibrium strategies before and after the adoption of blockchain. In studying blockchain-driven financing strategies, most existing scholars have adopted case studies or qualitative analyses; this study introduces the Stackelberg game model and considers both non-cooperative and cooperative games to explore the financing equilibrium strategies where logistics providers offer guarantees. Although some research has combined blockchain technology with the risk management of suppliers or financial institutions, and scholars have discussed the role of logistics providers in SCF, including their guarantees and profit distribution, very few papers have incorporated the risk preferences of logistics providers as on-chain guarantors and participants into systematic game analyses. This study, from the perspective of managerial behavior, considers the risk-averse behavior of enterprises and investigates its impact on financing strategies. Current research has paid attention to how blockchain reduces trust costs and improves transparency, but in multi-party cooperation scenarios, contract designs addressing how to fairly distribute the additional value generated on the chain and how to reasonably allocate the costs of blockchain construction and maintenance remain scarce. This paper, for the first time, explores, from the perspectives of game theory and contract design, how to design revenue-sharing and cost-sharing mechanisms under the background where blockchain nodes are jointly maintained by suppliers, logistics providers, e-commerce platforms, and financial institutions, and provides a detailed argumentation of their stability and incentive compatibility. Supply chain coordination strategies are commonly used in supply chain management, but few scholars have introduced revenue-sharing contracts and cost-sharing contracts when studying SCF strategies. Such contracts can effectively coordinate the interest conflicts of financing enterprises, and it is necessary to consider coordination strategies within the research.

3 Problem description and research hypothesis

3.1 Problem description

In an e-commerce supply chain composed of suppliers, logistics providers, and e-commerce platforms, suppliers face funding constraints, and traditional bank loans cannot fully meet their financial needs. Most suppliers are SMEs with weak credit foundations, resulting in limited loan amounts when applying to financial institutions. As an important member of the e-commerce supply chain, logistics providers can offer transportation services and financial services. Therefore, suppliers can seek guarantees from logistics providers to borrow from financial institutions, increase their financing amounts, and alleviate their own funding constraints. Considering the hierarchical nature of enterprise decision-making processes in financing and the characteristics of information asymmetry and strategic interdependence among parties, this paper selects the Stackelberg game model to depict the mutual influences among parties in decisions such as guarantee fees, financing interest rates, and product pricing. Using game analysis to build a financing model and explore the optimal financing strategy. There is a clear hierarchical relationship among e-commerce supply chain enterprises: suppliers first determine their financing needs and product wholesale prices; logistics providers then provide guarantees based on the suppliers' credit and mandates and are responsible for transporting products; e-commerce platforms determine retail prices based on wholesale prices and logistics costs; and banks decide loan interest rates and amounts. The Stackelberg game model can effectively depict the strategic interactions, the leader-follower relationship in the financing scenario, and is highly consistent with the hierarchical setting of the Stackelberg game model.

E-commerce suppliers are focusing on risk aversion. Due to uncertainties in the financing process such as financing interest rates, the value of collateral, and market selling prices, participants need to balance the pursuit of maximum returns with the avoidance of return volatility. This study selects the mean - standard deviation model to express the risk-averse of all participants. The utility of participants depends not only on expected returns but is also affected by return fluctuations. The mean - standard deviation method intuitively depicts the trade-off between pursuing high returns and avoiding volatility risks by combining expected returns with the standard deviation of returns into a weighted utility function. This method is widely adopted in supply chain finance and investment fields and can condense complex risk characteristics into a single indicator, facilitating analysis and solution. Compared with other risk measurement models, the mean - standard deviation can

directly use the variance of returns as a risk measure while retaining second-order information, improving the clarity of the model.

Blockchain technology has significant applications in the field of guarantee financing decisions, effectively improving the transparency of financing information and optimizing financing strategies [36]. The consortium blockchain architecture in blockchain technology can sufficiently protect the legitimate interests of all participating entities, safeguard data privacy, and improve transaction efficiency and governance controllability. Based on the consensus mechanism of blockchain, participants can read, write, and verify transaction information only under authorization, achieving controlled participation, data isolation, and high-performance consensus. The consortium blockchain nodes are pre-designated suppliers, logistics providers, e-commerce platforms, and banks. The maintenance costs mainly include hardware, operations, and smart contract development, without the need to purchase external Gas, and with minimal external attack risks. Costs can be quantified as relatively stable enterprise-level budgets, facilitating subsequent analyses of cost sensitivity in the model. Decision-making entities are concentrated at known nodes, and operations such as protocol upgrades, node additions, and deletions can be executed quickly through multi-party consensus, without the lengthy community discussions and voting required by public blockchains and avoiding risks. Since all nodes are core participants, the financing demands, logistics inspections, credit ratings, and other information shared on the chain have inherent credibility and do not require third-party verification, increasing trust among enterprises [37].

Under blockchain, to ensure the accuracy of data on the chain, it is usually necessary for multiple parties to collaboratively build a credible data chain [38]. In production or shipping stages, inventory change data is instantly synchronized to edge servers through integration with traditional ERP/WMS systems. In warehousing and logistics stages, Internet of Things (IoT) sensing devices are used to conduct real-time inventory counts, and collected weight, volume, and timestamp information for entry and exit from warehouses is uploaded to local gateways. Blockchain nodes are jointly maintained by suppliers, logistics providers, and core e-commerce/financial institutions. All nodes run lightweight clients, accept signed data packages sent by edge gateways, verify the digital signatures of producers/warehouse managers, and then package, consensus, and write inventory status to the blockchain. Based on smart contracts and consensus mechanisms, the authenticity and reliability of product transaction information are verified. If physical goods are tampered with, switched, or miscounted, the chain records will be inaccurate, leading to financial risks. Therefore, dual, dynamic cross-verification of product data on-chain and off-chain is needed to reduce financing risks.

Smart contracts automatically execute terms on the blockchain network, but to make them legally binding as contracts, all parties must explicitly agree to the contract terms through electronic signatures. Smart contract code must correspond to the hash of on-chain transactions, and contract terms must be saved either off-chain or on-chain in a human-readable form to meet subsequent judicial review needs. Smart contracts should clearly specify the conditions under which they automatically execute and send on-chain notifications to all parties before and after execution to ensure full awareness. Once deployed, smart contracts are executed automatically; if assets are incorrectly released or frozen due to technical vulnerabilities or hacker attacks, emergency arbitration or manual intervention mechanisms must be reserved. Given that the smart contract code itself is a legal document, any modifications to the code must go through on-chain upgrade proposals. Although blockchain inherently possesses anti-tampering and traceability features, during disputes, original off-chain sensor data or manual signature records should also be synchronously archived as evidence.

The financing and operation processes before and after the blockchain are shown in Figure 1.

Under the N mode without adopting blockchain, the credit cannot be transmitted. When suppliers apply for loans from internet financial institutions, they need third-party logistics companies to provide guarantees and must pay certain information verification costs. In financing and operation, suppliers produce products at a cost of c_s and sell them to downstream enterprises at a price of w_s^N. Without sufficient funds for production, suppliers apply for a guarantee from logistics companies and pay factoring fees of $c_s D^N \varepsilon^N$. They apply for loans from internet financial institutions at an interest rate of τ, and $\tau > \varepsilon^N$. To obtain better financing, they need to pay information verification costs of $c_s \delta D^N$. Logistics companies are responsible for providing transportation and financing services, receiving transportation income of w_d^N and incurring transportation costs of c_d. They charge guarantee return income of $c_s D^N \varepsilon^N$ for providing guarantees to suppliers. E-commerce platforms purchase products and pay transportation fees, then sell them to consumers at p^N. Risk-averse firms receive fund inflows, and suppliers promptly repay the funds. If suppliers default and fail to repay, logistics companies need to pay the unreturned loans to internet financial institutions.

Under the B mode driven by blockchain technology, e-commerce supply chains upload relevant transaction information to the blockchain during operations, and enterprises and financial institutions with permission can access this information. Blockchain's encryption algorithms, and other features help promote trust transmission among firms. Internet financial institutions can effectively evaluate the credit situation of capital-constrained suppliers based on data on the blockchain and blockchain trustworthiness without charging related information verification costs. In financing and operation, suppliers sell products with a production cost of c_s to e-commerce platforms at a sales price of w_s^B. Suppliers need to bear blockchain costs of $b_s^B D^B$ when introducing blockchain technology. Facing capital constraints, risk-averse suppliers apply for a guarantee from logistics companies and pay guarantee fees of $c_s \varepsilon^B D^B$. Internet financial institutions use data on the blockchain to assess the credit status of suppliers and logistics companies, provide loans to suppliers at an interest rate of τ. Logistics companies provide transportation services with transportation costs of c_d for logistics companies and charge transportation fees of w_d^N. Logistics companies provide financing guarantees for suppliers and charge factoring fees at a factoring return rate of ε^B. Logistics companies need to bear blockchain costs of $b_d^B D^B$ when introducing blockchain technology. E-commerce pays the price w_s^B to suppliers, pays transportation fees of w_d^B to logistics companies, pays blockchain costs of $b_e^B D^B$, and then sells at p^B.

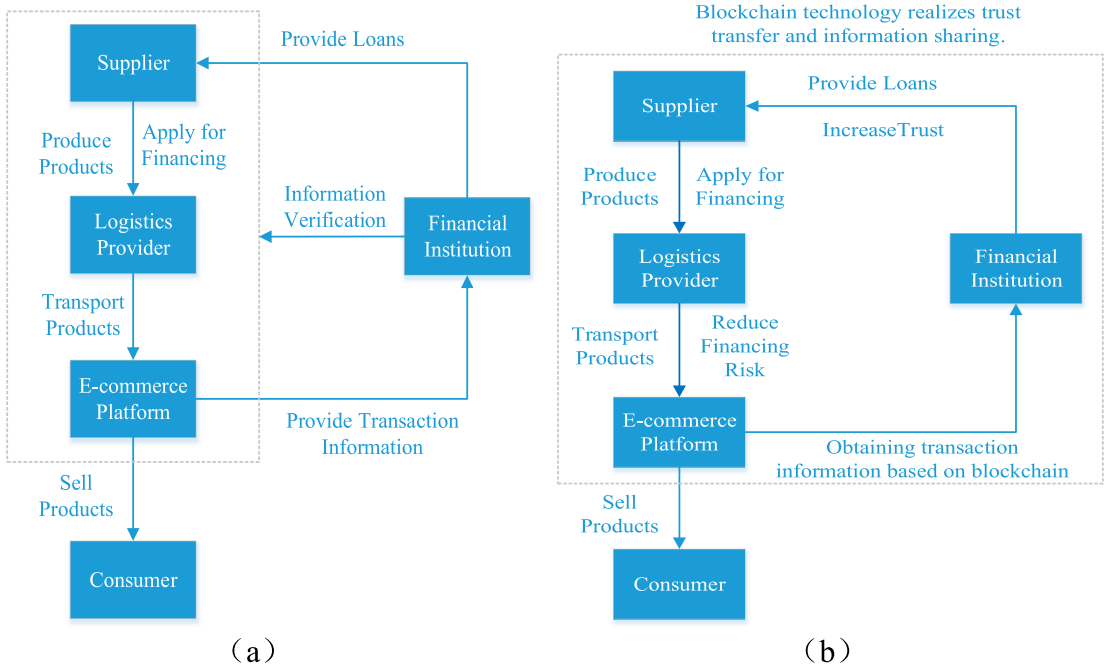


FIGURE 1
Decision order. **(a)** N Mode. **(b)** B Mode.

In a chain composed of vendors, logistics companies, and e-commerce, and guarantee financing model is constructed. Capital-constrained suppliers apply for guarantee financing, guaranteed by logistics companies, with internet financial institutions providing loans. The risk-averse behavior of managers is measured using the Conditional Value at Risk (CVaR) method, and a Stackelberg game model is built to analyze the guarantee financing strategies considering risk aversion before and after adopting blockchain. We also compare and analyze the non-cooperative game model where members make separate determination and the cooperative game model where members make centralized determination. To coordinate the conflicts between various members, revenue-sharing and risk-sharing are used to harmonize the firms.

The definitions of symbols are in Table 1.

TABLE 1 Definition of symbols.

Symbol	Meaning
D	Market demand function
ξ	Information collection efforts
a	Trust of product's authenticity
ε	Factoring return rate
δ	Unit cost of information verification
p	Retail prices of e-commerce enterprises
γ	Default probability of suppliers
η	Risk aversion coefficient
n	Number of purchasers
c	Cost of production
τ	Loan interest rate of internet financial institutions
b	Blockchain cost
w	Wholesale price of products
v	Consumer's valuation of products
θ	Cost sharing ratio

3.2 Research hypothesis

The research hypotheses proposed are:

(1) Reference scholars Tian et al. [39] and Li et al. [40], assume the market random function is $D = \tilde{n}\int_{p-a\xi}^{1} f(v)dv = \tilde{n}(1-p+a\xi)$. Here, $\tilde{n}$ is the consumer's number. v is the product valuation, and $f(v)$ is a function of v, which follows a distribution over [0,1]. Market orders are equivalent to market demand, and product market demand is directly proportional to a and ξ, and inversely proportional to p. Under the N mode, the trust level satisfies $0 < a < 1$. Under the B mode, the trust level satisfies $a = 1$.

(2) The firms are all risk averse. Financial institutions are risk-neutral and have ample funds. Reference scholars Choi et al. [41] and Gupta and Ivanov [42], the mean-standard is used to represent risk-averse behavior, with the formula: $U_i = E(\pi_i) - \eta_i\sqrt{VAR(\pi_i)}$. Here, $VAR(\pi_i) = E[\pi_i - E(\pi_i)]^2$. $E(\pi_i)$ is the expected financing return of enterprise i, $VAR(\pi_i)$ represents the variance of the random profit of firm i. η_i is the risk aversion, and the larger it, the higher the risk aversion.

4 Non-cooperative game financing equilibrium before and after the adoption of blockchain

In the non-cooperative game, enterprises tend to avoid risks. Considering the N mode without blockchain and the B mode with blockchain, we explore the financing equilibrium in different scenarios.

4.1 Non-cooperative game in mode N

The goods produced by the supplier are delivered to e-commerce by the logistics company, and the e-commerce purchases and sells them. The vendor faces a capital dilemma and subscribes to financing. The logistics company provides a guarantee and charges the corresponding guaranteed fee. Internet financial institutions provide loans and charge interest fees. During the financing process, it is necessary to verify the transaction information and the credit level of the vendor, and the vendor needs to bear the corresponding information verification costs. If the vendor does not repay the loan on time at the end of the period (with a probability γ), the logistics company needs to pay the unpaid amount. The financing profit function is shown in Equations 1–4:

$$\max_{w_s^N} \pi_s^N = \left[w_s^N - c_s^N\left(1 + \tau^N + \varepsilon^N + \delta\right)\right]D^N \tag{1}$$

$$\max_{w_d^N, \varepsilon^N} \pi_d^N = \left[w_d^N - c_d^N + c_s^N\varepsilon^N - \gamma c_s^N(1+\tau)\right]D^N \tag{2}$$

$$\max_{p^N} \pi_e^N = \left(p^N - w_s^N - w_d^N\right)D^N \tag{3}$$

$$\max_{\tau^N} \pi_b^N = c_s^N\tau^N D^N \tag{4}$$

where $\overline{D}^N = n\left(1 - p^N + a\xi + \epsilon\right)$, and the mean and standard deviation of the random variable ϵ are 0 and σ, respectively.

Following the approach of Choi et al. [41] and Gupta and Ivanov [42], the mean-standard method is used to represent the risk-averse, and the function is expressed as: $U\left(\pi_i^N\right) = E\left(\pi_i^N\right) - \eta_i\sqrt{var\left(\pi_i^N\right)}$. Based on the financing profit function of the enterprises, the expected profits are respectively: $E\left(\pi_s^N\right) = \left[w_s^N - c_s^N\left(1 + \tau + \varepsilon^N + \delta\right)\right]n\left(1 - p^N + a\xi\right)$, $E\left(\pi_d^N\right) = \left(w_d^N - c_d^N + c_s^N\varepsilon^N - \gamma c_s^N(1+\tau)\right)n\left(1 - p^N + a\xi\right)$, $E\left(\pi_e^N\right) = \left(p^N - w_s^N - w_d^N\right)n\left(1 - p^N + a\xi\right)$. The variances are respectively: $var\left(\pi_s^N\right) = \left[w_s^N - c_s^N\left(1 + \tau + \varepsilon^N + \delta\right)\right]^2 n^2$, $var\left(\pi_d^N\right) = \left(w_d^N - c_d^N + c_s^N\varepsilon^N - \gamma c_s^N(1+\tau)\right)^2 n^2$, $var\left(\pi_e^N\right) = \left(p^N - w_s^N - w_d^N\right)^2 n^2$. Substituting these into the mean - standard deviation function, we can obtain: $U\left(\pi_s^N\right) = \left[w_s^N - c_s^N\left(1 + \tau + \varepsilon^N + \delta\right)\right]n\left(1 - p^N + a\xi - \eta\sigma\right)$, $U\left(\pi_d^N\right) = \left(w_d^N - c_d^N + c_s^N\varepsilon^N - \gamma c_s^N(1+\tau)\right)n\left(1 - p^N + a\xi - \eta\sigma\right)$, $U\left(\pi_e^N\right) = \left(p^N - w_s^N - w_d^N\right)n\left(1 - p^N + a\xi - \eta\sigma\right)$. By solving the second-order partial derivatives of the formula, it is found that the function is concave.

Theorem 1: Using Stackelberg game, the best equilibrium for each variable can be obtained respectively. The supplier's best wholesale price is $w_s^{N^*} = \frac{a\xi + \eta\sigma - c_d - c_s\gamma(\tau+1) + c_s(\delta + 2\varepsilon + \tau + 1) + 1}{2}$. The logistics provider's optimal logistics fee is $w_d^{N^*} = \frac{a\xi - \eta\sigma - w_s + c_d - c_s\varepsilon + c_s\gamma(\tau+1) + 1}{2}$. The optimal guarantee fee is $\varepsilon^{N^*} = \frac{c_d + c_s\gamma(\tau+1) - a\xi + \eta\sigma + w_s - 1}{c_s}$. The e-commerce platform's best retail price is $p_e^{N^*} = \frac{(w_s + w_d) + (a\xi - \eta\sigma + 1)}{2}$. The financial institution's best interest rate is $\tau^{N^*} = \frac{a\xi - \eta\sigma + 1 - w_s - c_d + c_s\varepsilon - c_s\gamma}{2c_s\gamma}$.

Substituting the above top solutions into the profit functions, and after rearranging, the top financing profits for the supplier, logistics company, e-commerce and financial institutions can be shown in Equations 5–8:

$$\pi_s^{N*} = \frac{n\left(c_s + c_d + c_s\delta + c_s\gamma - a\xi + c_s\tau + \eta\sigma + c_s\gamma\tau - 1\right)^2}{16} \tag{5}$$

$$\pi_d^{N*} = \frac{n\left(c_s - 2c_d - 3w_s + c_s\delta + 3c_s\varepsilon - 2c_s\gamma + 2a\xi + c_s\tau - 2\eta\sigma - 3c_s\gamma\tau + 2\right)^2}{32} \tag{6}$$

$$\pi_e^{N*} = \frac{n\left(c_d + w_s - c_s\varepsilon + c_s\gamma - a\xi + \eta\sigma + c_s\gamma\tau - 1\right)\left(2c_s - c_d - 3w_s + 2c_s\delta + 3c_s\varepsilon - c_s\gamma + a\xi + 2c_s\tau - \eta\sigma - c_s\gamma\tau + 1\right)}{4} \tag{7}$$

$$\pi_b^{N*} = \frac{n\left(c_d + w_s - c_s\varepsilon + c_s\gamma - a\xi + \eta\sigma - 1\right)\left(c_s - 2c_d - 3w_s + c_s\delta + 3c_s\varepsilon - 2c_s\gamma + 2a\xi + c_s\tau - 2\eta\sigma - 3c_s\gamma\tau + 2\right)}{16\gamma} \tag{8}$$

4.2 Non-cooperative games in mode B

After the introduction of blockchain, although the blockchain cost (node deployment and operation and maintenance costs under the alliance chain architecture) is increased, the sharing degree of transaction information is improved and the trust transfer between enterprises is promoted. Based on the technical characteristics of instant consensus and reliable nodes in the alliance chain in the blockchain, each participant does not need to bear the extra cost of information verification under the chain and does not need to worry about the delay of transaction confirmation on the chain. Logistics providers and internet banks are more willing to provide finance services for suppliers. In the financing process, suppliers pay guarantee fees to logistics companies and interest fees to internet financial institutions. All members need to pay blockchain costs, and the finance profit function is shown in Equations 9–12:

$$\max_{w_s^B} \pi_s^B = \left[w_s^B - c_s^B\left(1 + \tau^B + \varepsilon^B\right) - b\right]D^B \tag{9}$$

$$\max_{w_d^B, \varepsilon^B} \pi_d^B = \left(w_d^B - c_d^B + c_s^B\varepsilon^B - b\right)D^B \tag{10}$$

$$\max_{p^N} \pi_e^B = \left(p^B - w_s^B - w_d^B - b\right)D^B \tag{11}$$

$$\max_{\tau^B} \pi_b^B = c_s^B \tau^B D^B \tag{12}$$

where $\overline{D}^B = n\left(1 - p^B + a\xi + \epsilon\right)$.

According to the utility function, the expected profits of the enterprises are respectively: $E\left(\pi_s^B\right) = \left[w_s^B - c_s^B\left(1+\tau+\varepsilon^B\right) - b\right]n\left(1 - p^B + a\xi\right)$, $E\left(\pi_d^B\right) = \left(w_d^B - c_d^B + c_s\varepsilon^B - b\right)n\left(1 - p^B + a\xi\right)$, $E\left(\pi_e^B\right) = \left(p^B - w_s^B - w_d^B - b\right)n\left(1 - p^N + a\xi\right)$. The variances are respectively: $var\left(\pi_s^B\right) = \left[w_s^B - c_s^B\left(1+\tau+\varepsilon^B\right) - b\right]^2 n^2$, $var\left(\pi_d^B\right) = \left(w_d^B - c_d^B + c_s\varepsilon^B - b\right)^2 n^2$, $var\left(\pi_e^B\right) = \left(p^B - w_s^B - w_d^B - b\right)^2 n^2$. Substituting these into the mean - standard deviation function $U\left(\pi_s^B\right) = \left[w_s^B - c_s^B\left(1+\tau+\varepsilon^B\right) - b\right]n\left(1 - p^B + a\xi - \eta\sigma\right)$, $U\left(\pi_d^B\right) = \left(w_d^B - c_d^B + c_s\varepsilon^B - b\right)n\left(1 - p^B + a\xi - \eta\sigma\right)$, $U\left(\pi_e^B\right) = \left(p^B - w_s^B - w_d^B - b\right)n\left(1 - p^B + a\xi - \eta\sigma\right)$. By solving the second-order partial derivatives, it is found that the function is concave and has a top solution.

Theorem 2: According to Stackelberg, the top equilibrium for each variable can be obtained respectively. The supplier's best wholesale price is $w_s^{B*} = \frac{b + c_s(\varepsilon+\tau+1) - w_d + (a\xi - \eta\sigma + 1)}{2}$. The logistics provider's optimal logistics fee is $w_d^{B*} = a\xi - \eta\sigma - w_s - b + 1$. The optimal guarantee fee is $\varepsilon^{B*} = \frac{2b + c_d + w_s - a\xi + \eta\sigma - 1}{c_s}$. The e-commerce platform's best retail price is $p_e^{B*} = \frac{(b + w_s + w_d) + (a\xi - \eta\sigma + 1)}{2}$. The financial institution's optimal interest rate is $\tau^{B*} = \frac{2w_s + w_d - c_s(\varepsilon+1) - (a\xi - \eta\sigma + 1)}{2c_1}$.

Substituting the top solutions into the expected profit functions, the best finance profits for enterprises and financial institutions can be obtained in Equations 13–16 respectively:

$$\pi_s^{B*} = \frac{n\left(3b + c_s + c_d - a\xi + c_s\tau + \eta\sigma - 1\right)^2}{16} \tag{13}$$

$$\pi_d^{B*} = \frac{n\left(w_s - b - c_s - c_s\varepsilon - c_s\tau\right)\left(c_d - 2b + w_s - c_s\varepsilon - a\xi + \eta\sigma - 1\right)}{8} \tag{14}$$

$$\pi_e^{B*} = \frac{n\left(2b + c_d + w_s - c_s\varepsilon - a\xi + \eta\sigma - 1\right)\left(2c_s - c_d - 3w_s + 3c_s\varepsilon + a\xi + 2c_s\tau - \eta\sigma + 1\right)}{2} \tag{15}$$

$$\pi_b^{B*} = \frac{n\left(2b + c_d + w_s - c_s\varepsilon - a\xi + \eta\sigma - 1\right)\left(b + c_s - 2c_d - 3w_s + 3c_s\varepsilon + 2a\xi + c_s\tau - 2\eta\sigma + 2\right)}{16} \tag{16}$$

4.3 Comparative analysis

By analyzing the top equilibrium solutions of various decision variables and the top financing profit functions of various decision-making entities, this paper explores the conditions and constraints for enterprises to introduce blockchain from the financing income. Before and after the blockchain, the changes in the intensity of various decision variables of risk-averse members are compared and analyzed, and the effect of the risk aversion is investigated.

Proposition 1: When the decision variables satisfy certain conditions $b > \frac{c_s\delta + c_s\gamma + c_s\gamma\tau}{3}$, supply chain members can obtain more benefits under blockchain than without it. That is, $\overline{\pi}_s^{N^*} < \overline{\pi}_s^{B^*}$, $\overline{\pi}_b^{N^*} < \overline{\pi}_b^{B^*}$.

Proof: Before blockchain, the top financing profit of the vendor is $\pi_s^{N*} = \frac{n(c_s + c_d + c_s\delta + c_s\gamma - a\xi + c_s\tau + \eta\sigma + c_s\gamma\tau - 1)^2}{16}$, and after blockchain, the finance profit is $\pi_s^{B*} = \frac{n(3b + c_s + c_d - a\xi + c_s\tau + \eta\sigma - 1)^2}{16}$. When certain conditions $b > \frac{c_s\delta + c_s\gamma + c_s\gamma\tau}{3}$ are met, $\left(c_s + c_d + c_s\delta + c_s\gamma - a\xi + c_s\tau + \eta\sigma + c_s\gamma\tau - 1\right)^2 < \left(3b + c_s + c_d - a\xi + c_s\tau + \eta\sigma - 1\right)^2$, i.e., $\pi_s^{N*} < \pi_s^{B*}$. Similarly, it can be proven that $\pi_b^{N*} < \pi_b^{B*}$.

The cost of blockchain technology is a critical factor in whether supply chain members adopt blockchain; both capital-constrained suppliers and well-funded financial institutions focus on maximizing their own interests. When the condition is met $b > \frac{c_s\delta + c_s\gamma + c_s\gamma\tau}{3}$, adopting blockchain can create more financing benefits, and considering the impact of returns will further incline towards the adoption of blockchain technology, thereby helping to achieve a win-win situation between e-commerce supply chains and financial institutions. As the effort in information collection decreases, the threshold range will expand further, making enterprises more dependent on blockchain technology, and the benefits blockchain brings to supply chain enterprises will be greater. When the cost of blockchain technology meets the conditions, blockchain not only facilitates information sharing and brings more financing benefits but also helps in the transmission of trust and improvement of operational efficiency, enhancing overall profit levels. The introduction of blockchain can effectively protect the interests of logistics providers; the automatic execution of smart contracts can help prevent supplier defaults, reduce financing risks, and increase financing returns. Immutable records and real-time monitoring lower default risks and reduce expensive risk premiums. Blockchain makes information more transparent, reduces fraud risks, increases the likelihood of repayment, and lowers the probability of supplier defaults. By coordinating incentives through trust enhancement and risk reduction, the adoption of blockchain creates a win-win situation for suppliers and financial institutions. The condition for profitability lies in balancing the cost of blockchain with the actual benefits derived from increased transparency and efficiency.

Proposition 2: When $b < w_d - c_d - c_s\gamma(\tau+1) + c_s(\delta+\varepsilon)$, the optimal sales prices in the e-commerce supply chain satisfy the following relationship between scenarios without and with blockchain technology: $w_s^{B*} < w_s^{N*}$, $p_e^{B*} > p_e^{N*}$.

Proof: Based on the optimal solutions of each decision variable, the supplier's optimal wholesale prices before and after blockchain adoption are $w_s^{N*} = \frac{a\xi + \eta\sigma - c_d - c_s\gamma(\tau+1) + c_s(\delta + 2\varepsilon + \tau + 1) + 1}{2}$ and $w_s^{B*} = \frac{b + c_s(\varepsilon+\tau+1) - w_d + (a\xi - \eta\sigma + 1)}{2}$ respectively. $w_s^{B*} - w_s^{N*} = \frac{b + c_s(\varepsilon+\tau+1) - w_d + (a\xi - \eta\sigma + 1)}{2} - \frac{a\xi + \eta\sigma - c_d - c_s\gamma(\tau+1) + c_s(\delta + 2\varepsilon + \tau + 1) + 1}{2}$, When $b < w_d - c_d - c_s\gamma(\tau+1) + c_s(\delta+\varepsilon)$, $w_s^{B*} - w_s^{N*} < 0$, clearly establishing the relationship $w_s^{B*} < w_s^{N*}$. The e-commerce platform's optimal retail prices are $p_e^{N*} = \frac{(w_s + w_d) + (a\xi - \eta\sigma + 1)}{2}$ and $p_e^{B*} = \frac{(b + w_s + w_d) + (a\xi - \eta\sigma + 1)}{2}$ respectively, thus $p_e^{B*} > p_e^{N*}$.

When blockchain technology satisfies condition $b < w_d - c_d - c_s\gamma(\tau+1) + c_s(\delta+\varepsilon)$, and the blockchain benchmark cost is below the threshold, the application of blockchain can significantly optimize the cost structure and risk allocation of the supply chain, thereby changing pricing strategies. Through comparative analysis of the changes in various decision variables of e-commerce supply chain members before and after adopting blockchain, it is found that the

adoption of blockchain can raise trust transmission among supply chain members, reduce information verification costs between members, lower the risk premium of suppliers, replace manual reviews with smart contracts, reduce transaction frictions, lower operating costs, and lead to a decrease in sales prices after adopting blockchain. However, the adoption of blockchain will also increase technical investment costs, especially having a significant impact on e-commerce platforms, and under blockchain, the final retail price will be higher. But the adoption of blockchain can provide consumers with more transparent information, enabling consumers to trace product information, reassure them about product quality, attract more consumers to purchase products, and increase market demand for the supply chain. The traceability and anti-counterfeiting capabilities provided by blockchain make consumers more willing to pay a premium for products. Retailers can build a transparent supply chain brand image through blockchain technology, thereby increasing product added value. The initial cost of implementing blockchain can be partially passed on to consumers through higher retail prices, optimizing the cost structure. Blockchain achieves the reconstruction of the balance between supply chain costs and value by optimizing supply chain efficiency and creating additional value, enhancing consumer trust.

Proposition 3: The effect of risk aversion coefficient on guaranteed rates under both non-blockchain and blockchain scenarios is as follows: $\varepsilon^{N^*}, \varepsilon^{B^*}$ is positively correlated with η.

In an environment of uncertain market demand, a company's risk attitude has a significant influence on finance decisions. The guaranteed interest rate of logistics providers increases with the rise in risk aversion. Under a higher risk aversion attitude, logistics providers, to ensure their own level of returns, will set higher guarantee interest rates. In traditional financing models, the guaranteed interest rate is jointly determined by the value of collateral, liquidity, and market risk. A high-risk aversion coefficient typically manifests as a conservative financial strategy, leading to a rise in the guaranteed interest rate as the risk aversion coefficient rises. After blockchain, blockchain, through the transparency of collateral information, it reduces information asymmetry and decreases information verification costs. Although technology reduces operational risks, logistics providers with high risk aversion still require stringent guarantee terms to enhance trust between enterprises. Blockchain does not eliminate market risks, and the guaranteed interest rate still reflects the risk preference of logistics providers, thus maintaining a positive correlation with the risk aversion. Under the traditional model, the negative correlation between the financing interest rate and the risk aversion means that suppliers with a high-risk aversion tend to adopt conservative operational strategies and pursue stable cash flows. For financial institutions, high-risk-averse suppliers, due to their lower probability of default, are considered better credit customers. Financial institutions reward high-credit customers with lower interest rates to attract low-risk, capital-constrained suppliers and optimize their risk asset portfolios. Under the blockchain model, the positive correlation between the finance interest rate and the risk aversion coefficient indicates that blockchain changes the risk assessment mechanism, and transparent data allows financial institutions to accurately evaluate the true risk status of borrowers. Conservative strategies under risk aversion may reflect business stability, but excessive risk aversion may suppress growth potential. For financial institutions, high-risk-averse suppliers may be interpreted as lacking growth momentum or insufficient risk resistance capabilities. In a blockchain environment, financial institutions prefer suppliers with moderate risk-taking and therefore charge higher interest rates to overly risk-averse suppliers to compensate for potential low growth expectations.

5 Cooperative game financing equilibrium after the adoption of blockchain

When blockchain is low, members can achieve higher financing returns. Blockchain technology enhances trust among vendors, logistics, and e-commerce, forming an alliance through information sharing. When e-commerce supply chain members cooperate, a cooperative game model under blockchain is constructed to compare and analyze the financing equilibrium strategies. Partial cooperation and collective cooperation among members are conducive to improving the overall financing returns. Members write information into blockchain technology can regulate the decision-making behaviors of enterprises, improving transaction quality and financing efficiency. In the cooperative game, different cooperation modes are considered separately, including cooperation between suppliers and logistics, cooperation between logistics and e-commerce, and tripartite cooperation among suppliers, logistics, and platforms, to explore the financing equilibrium decisions of risk-averse firms under different modes.

5.1 Cooperative game of alliance between supplier and logistics company

When the supplier and logistics company form an alliance, the supplier and logistics company are the leaders, and the platforms are the followers. The financing profit of members is shown in Equations 17–19:

$$\max_{w_d^B} \pi_{sd}^{BC1} = \left[w_d^{BC1} - c_s^{BC1}\left(1+\tau^{BC1}\right) - c_d^{BC1} - 2b^{BC1}\right] D^{BC1} \tag{17}$$

$$\max_{p^B} \pi_e^{BC1} = \left(p^{BC1} - w_d^{BC1} - b^{BC1}\right) D^{BC1} \tag{18}$$

$$\max_{\tau^{BC}} \pi_b^{BC1} = c_s^{BC1} \tau^{BC1} D^{BC1} \tag{19}$$

According to the financing profit, the expected profits are respectively: $E\left(\pi_{sd}^{BC1}\right) = \left[w_d^{BC1} - c_s^{BC1}\left(1+\tau^{BC1}\right) - c_d^{BC1} - 2b^{BC1}\right] n(1 - p^{BC1} + a\xi)$, $E\left(\pi_e^{BC1}\right) = \left(p^{BC1} - w_d^{BC1} - b^{BC1}\right) n\left(1 - p^{BC1} + a\xi\right)$. The variances are respectively: $var\left(\pi_{sd}^{BC1}\right) = \left[w_d^{BC1} - c_s^{BC1}\left(1+\tau^{BC1}\right) - c_d^{BC1} - 2b^{BC1}\right]^2 n^2, var\left(\pi_e^{BC1}\right) = \left(p^{BC1} - w_d^{BC1} - b^{BC1}\right)^2 n^2$. Substituting these into the mean - standard deviation function $U\left(\pi_i^{BC}\right) = E\left(\pi_i^{BC}\right) - \eta_i \sqrt{var\left(\pi_i^{BC}\right)}$, we can obtain: $U\left(\pi_{sd}^{BC1}\right) = \left[w_d^{BC1} - c_s^{BC1}\left(1+\tau^{BC1}\right) - c_d^{BC1} - 2b^{BC1}\right] n\left(1 - p^{BC1} + a\xi - \eta\sigma\right)$, $U\left(\pi_e^{BC1}\right) = \left(p^{BC1} - w_d^{BC1} - b^{BC1}\right) n\left(1 - p^{BC1} + a\xi - \eta\sigma\right)$. By solving the second-order partial derivatives, it is found that the function is concave.

Theorem 3: According to Stackelberg, the top equilibrium solutions for each variable can be obtained respectively When suppliers

and logistics providers cooperate, the optimal product sales price is $w_{sd}^{BC1^*} = \frac{b+c_d+c_s(\tau+1)+a\xi-\eta\sigma+1}{2}$, and the best retail price of e-commerce is $p_e^{BC1^*} = \frac{b+w_d+a\xi-\eta\sigma+1}{2}$. The best loan interest rate of financial institutions is $\tau^{BC1^*} = \frac{a\xi-\eta\sigma+1-3b-c_s-c_d}{2c_s}$.

Substituting the top solutions into the financing gain function, the top profits for each entity can be obtained in Equations 20–22 respectively:

$$\pi_{sd}^{BC1*} = \frac{n(3b+c_s+c_d-a\xi+c_s\tau+\eta\sigma-1)^2}{8} \tag{20}$$

$$\pi_e^{BC1*} = \frac{n(2b+c_s+c_d-w_d+c_s\tau)(b+w_d-a\xi+\eta\sigma-1)}{4} \tag{21}$$

$$\pi_b^{BC1*} = \frac{c_s\eta\tau(a\xi-\eta\sigma-b-w_d+1)}{2} \tag{22}$$

5.2 Cooperative game of alliance between logistics company and E-commerce platform

When the logistics company and e-commerce form an alliance, the supplier remains the leader, and the logistics and e-commerce are the followers. The financing function is shown in Equations 23–25:

$$\max_{w_s^B} \pi_s^{BC2} = \left[w_s^{BC2} - c_s^{BC2}\left(1+\tau^{BC2}+\varepsilon^{BC2}\right) - b^{BC2}\right]D^{BC2} \tag{23}$$

$$\max_{p^B} \pi_{de}^{BC2} = \left(p^{BC2} - w_s^{BC2} - c_d^{BC2} + c_s\varepsilon^{BC2} - 2b^{BC2}\right)D^{BC2} \tag{24}$$

$$\max_{\tau^{BC}} \pi_b^{BC2} = c_s^{BC2}\tau^{BC2}D^{BC2} \tag{25}$$

Using the mean-standard to rundown the risk attitude, the expected profits and variances are respectively: $E\left(\pi_s^{BC2}\right) = \left[w_s^{BC2} - c_s^{BC2}\left(1+\tau^{BC2}+\varepsilon^{BC2}\right) - b^{BC2}\right]n\left(1-p^{BC2}+a\xi\right)$, $E\left(\pi_{de}^{BC2}\right) = \left(p^{BC2} - w_s^{BC2} - c_d^{BC2} + c_s\varepsilon^{BC2} - 2b^{BC2}\right)n\left(1-p^{BC2}+a\xi\right)$, $var\left(\pi_s^{BC2}\right) = \left[w_s^{BC2} - c_s^{BC2}\left(1+\tau^{BC2}+\varepsilon^{BC2}\right) - b^{BC2}\right]^2 n^2$, $var\left(\pi_{de}^{BC2} = \left(p^{BC2} - w_s^{BC2}\right)\right.$ $\left.-c_d^{BC2} + c_s\varepsilon^{BC2} - 2b^{BC2}\right)^2 n^2$. Substituting these into the mean - standard deviation function $U\left(\pi_s^{BC2}\right) = \left[w_s^{BC2} - c_s^{BC2}\left(1+\tau^{BC2}+\varepsilon^{BC2}\right)\right.$ $\left.-b^{BC2}\right]n\left(1-p^{BC2}+a\xi-\eta\sigma\right)$, $U\left(\pi_{de}^{BC2}\right) = \left(p^{BC2} - w_s^{BC2} - c_d^{BC2} + c_s\varepsilon^{BC2}\right.$ $\left.-2b^{BC2}\right)n\left(1-p^{BC2}+a\xi-\eta\sigma\right)$. By solving the second-order partial derivatives, it is found that the function is concave.

Theorem 4: According to Stackelberg, the top equilibrium solutions for each variable can be obtained respectively. The optimal wholesale price of suppliers is $w_s^{BC2^*} = \frac{a\xi-\eta\sigma+1-c_d+b+c_s(2\varepsilon+\tau+1)}{2}$. The optimal retail price when logistics providers and e-commerce platforms cooperate is $p_{de}^{BC2^*} = \frac{2b+c_d+w_s-c_s\varepsilon+a\xi-\eta\sigma+1}{2}$. The optimal loan interest of financial institutions is $\tau^{BC2^*} = \frac{a\xi-\eta\sigma+1-b-c_s(\varepsilon+1)-(2b+c_d-c_s\varepsilon)}{2c_s}$.

Substituting the top solutions into the financing profit function, and after rearranging, we can obtain in Equations 26–28:

$$\pi_s^{BC2*} = \frac{n(3b+c_s+c_d-a\xi+c_s\tau+\eta\sigma-1)^2}{8} \tag{26}$$

$$\pi_{de}^{BC2*} = \frac{n(3b+c_s+c_d-a\xi+c_s\tau+\eta\sigma-1)^2}{16} \tag{27}$$

$$\pi_b^{BC2*} = \frac{c_s\eta\tau(a\xi-3b-c_s-c_d-c_s\tau-\eta\sigma+1)}{4} \tag{28}$$

5.3 Cooperative game of alliance among suppliers, logistics company and E-commerce platform

When the vendor, logistics company, and e-commerce form an alliance, the members act, and the top decision is to seek the maximization of overall benefits. According to the Stackelberg game, the financing profit function currently is shown in Equations 29, 30:

$$\max_{p^B} \pi_{sde}^{BC3} = \left[p^{BC3} - c_s^{BC3}\left(1+\tau^{BC3}\right) - c_d^{BC3} - 3b^{BC3}\right]D^{BC3} \tag{29}$$

$$\max_{\tau^B} \pi_b^{BC3} = c_s^{BC3}\tau^{BC3}D^{BC3} \tag{30}$$

The risk-averse expected financing profit and variance are respectively: $E\left(\pi_{se}^{BC3}\right) = \left[p^{BC3} - c_s^{BC3}(1+\tau) - c_d^{BC3} - 3b^{BC3}\right]n\left(1-p^{BC3}+a\xi\right)$, $var\left(\pi_{se}^{BC3}\right) = \left[p^{BC3} - c_s^{BC3}(1+\tau) - c_d^{BC3} - 3b^{BC3}\right]^2 n^2$. Substituting these into the mean - standard deviation function, we get: $U\left(\pi_{se}^{BC3}\right) = \left[p^{BC3} - c_s^{BC3}(1+\tau) - c_d^{BC3} - 3b^{BC3}\right]n\left(1-p^{BC3}+a\xi-\eta\sigma\right)$. Taking the second-order derivative of the formula, it is found that the function is concave.

Theorem 5: According to Stackelberg, the top equilibrium solution is: $p_{sde}^{BC3^*} = \frac{3b+c_d+c_s(\tau+1)+a\xi-\eta\sigma+1}{2}$, $\tau^{BC3^*} = \frac{(a\xi-\eta\sigma+1)-(3b+c_s+c_d)}{2c_s}$.

Substituting the top solution into the expected financing profit function, we obtain Equations 31, 32:

$$\pi_{sde}^{BC3*} = \frac{n(3b+c_s+c_d-a\xi+\eta\sigma-1)^2}{16} \tag{31}$$

$$\pi_b^{BC3*} = \frac{n(3b+c_s+c_d-a\xi+\eta\sigma-1)^2}{8} \tag{32}$$

5.4 Comparative analysis

After solving the model using the Stackelberg game, the top variables and top financing profits are get. A comparative analysis of the relationships between the prices, market demand, and financing profits under different alliances is conducted.

Proposition 4: When certain conditions are met $b < w_d + c_s\varepsilon - c_d - w_s$ and $w_s > \frac{w_d+2c_s\varepsilon-c_d+c_s\tau+c_s}{2}$, under blockchain, the retail prices of different alliances satisfy: $p_e^{BC1^*} > p_{de}^{BC2^*} > p_{sde}^{BC3^*}$.

Proof: The optimal retail prices under different cooperative game modes are $p_e^{BC1^*} = \frac{b+w_d+a\xi-\eta\sigma+1}{2}$, $p_{de}^{BC2^*} = \frac{2b+c_d+w_s-c_s\varepsilon+a\xi-\eta\sigma+1}{2}$, $p_{sde}^{BC3^*} = \frac{3b+c_d+c_s(\tau+1)+a\xi-\eta\sigma+1}{2}$, respectively. When $b < w_d + c_s\varepsilon - c_d - w_s$, that is $p_e^{BC1^*} - p_{de}^{BC2^*} = \frac{b+w_d+a\xi-\eta\sigma+1}{2} - \frac{2b+c_d+w_s-c_s\varepsilon+a\xi-\eta\sigma+1}{2} = \frac{-b+w_d-c_d-w_s+c_s\varepsilon}{2} > 0$. $p_e^{BC1^*} > p_{de}^{BC2^*}$, $p_{de}^{BC2^*} - p_{sde}^{BC3^*} = \frac{2b+c_d+w_s-c_s\varepsilon+a\xi-\eta\sigma+1}{2} - \frac{3b+c_d+c_s(\tau+1)+a\xi-\eta\sigma+1}{2} = \frac{w_s-c_s\varepsilon-b-c_s\tau-c_s}{2}$, because $b < w_d + c_s\varepsilon - c_d - w_s$, so $p_{de}^{BC2^*} - p_{sde}^{BC3^*} > \frac{2w_s-2c_s\varepsilon-c_s\tau-c_s-w_d+c_d}{2}$. Because $w_s > \frac{w_d+2c_s\varepsilon-c_d+c_s\tau+c_s}{2}$, therefore

$\frac{2w_s-2c_s\varepsilon-c_s\tau-c_s-w_d+c_d}{2} > 0$, that is $p_{de}^{BC2^*} > p_{sde}^{BC3^*}$, to sum up $p_e^{BC1^*} > p_{de}^{BC2^*} > p_{sde}^{BC3^*}$.

When the cost of blockchain and the supplier's selling price meet certain conditions, from the perspective of cost-profit, the more companies join the supply chain alliance, the more effectively they can reduce the intermediate operating costs through cooperation between enterprises. When pursuing profit maximization, e-commerce platforms, due to the reduction of intermediate costs between enterprises, are prompted to lower the price at which they sell to market consumers. Lower selling prices will attract more consumers to buy, leading to a rise in market demand. Greater market demand helps the e-commerce supply chain quickly recover funds, alleviating the supplier's financial constraints. The number of members and role divisions in different alliance structures directly affect the profit distribution model, which in turn shapes the retail pricing strategy. When suppliers and logistics providers form an alliance, the profit distribution is relatively concentrated. Suppliers transfer part of the costs by raising wholesale prices, while logistics providers reduce operating costs by optimizing logistics efficiency. With fewer members, the coordination costs within the alliance are low, and both parties may opt for value-based pricing, i.e., using blockchain technology to improve product quality or service reliability, resulting in higher retail prices. When logistics providers and e-commerce platforms form an alliance, the platform faces market competition pressure, forcing it to lower retail prices to attract consumers. E-commerce platforms rely on traffic and user scale to generate revenue and tend to lower per-transaction prices through subscription models or cost-plus pricing to expand market share. When all three parties form an alliance, there are the most members, and the profit distribution chain is the longest. Although blockchain technology improves overall efficiency, the cost-sharing mechanism leads to intensified competition for profits among the parties. E-commerce platforms may further compress retail prices. Additionally, suppliers and logistics providers, needing to share data and technical benefits with e-commerce platforms, may be forced to accept lower wholesale prices, indirectly lowering retail prices. Blockchain technology, by reducing information asymmetry and improving supply chain efficiency, reshapes the pricing logic of different alliances. When the cost of blockchain is below the threshold, the application of the technology is economically feasible. If the cost of blockchain is too high, the alliance may abandon the technology investment, causing retail prices to revert to traditional models. Higher wholesale prices reflect the dominant position of suppliers in the alliance. Blockchain technology may enhance suppliers' control over prices through dynamic pricing models. The effectiveness of blockchain technology applications is profoundly influenced by the alliance structure. The more members there are, the more complex the profit distribution becomes, and the marginal diminishing effect of technical benefits becomes more significant, ultimately leading to a decrease in retail prices as the alliance size expands.

Proposition 5: When $b > a\xi - \eta\sigma - 2w_d + 1 + c_s + c_d + c_s\tau$, the total income of financial institutions under the two-party cooperation is satisfied: $\pi_b^{BC1^*} < \pi_b^{BC2^*}$.

Proof: Under cooperation, the financial institution's financing profits are $\pi_b^{BC1^*} = \frac{c_s\eta\tau(a\xi-\eta\sigma-b-w_d+1)}{2}$ and $\pi_b^{BC2^*} = \frac{c_s\eta\tau(a\xi-3b-c_s-c_d-c_s\tau-\eta\sigma+1)}{4}$ respectively. $\pi_b^{BC1^*} - \pi_b^{BC2^*} = \frac{c_s\eta\tau(a\xi-\eta\sigma-b-w_d+1)}{2} - \frac{c_s\eta\tau(a\xi-3b-c_s-c_d-c_s\tau-\eta\sigma+1)}{4} = \frac{c_s\eta\tau[2(a\xi-\eta\sigma-b-w_d+1)-(a\xi-3b-c_s-c_d-c_s\tau-\eta\sigma+1)]}{4}$. Since $b > a\xi - \eta\sigma - 2w_d + 1 + c_s + c_d + c_s\tau$ and $a\xi - \eta\sigma + b - 2w_d + 1 + c_s + c_d + c_s\tau < 0$, then $\pi_b^{BC1^*} - \pi_b^{BC2^*} < 0$, i.e., $\pi_b^{BC1^*} < \pi_b^{BC2^*}$.

Comparative analysis of the overall gains of financial institutions under different alliances reveals that when the blockchain cost satisfies $b > a\xi - \eta\sigma - 2w_d + 1 + c_s + c_d + c_s\tau$, the financing returns of financial institutions under different alliances will vary. When the blockchain cost is within a certain threshold range, under the cooperation mode between logistics providers and e-commerce platforms, financial institutions will obtain higher financing returns through optimized resource allocation and risk sharing. Under the cooperation mode between suppliers and logistics providers, blockchain optimizes inventory management and reduces storage losses. Smart contracts execute automatically, lowering the cost of manual reviews. However, under this model, services are only provided to suppliers and logistics providers, lacking direct channels to consumers, making it difficult to leverage consumer trust premiums. Under the logistics provider-e-commerce platform cooperation model, the large user base of e-commerce platforms can significantly spread the blockchain costs. Integration of blockchain with e-commerce data enables precise demand forecasting and inventory optimization. The traceability of blockchain enhances product added value, allowing e-commerce platforms to charge premiums. The diversified businesses of e-commerce platforms can smooth out market demand fluctuations, and smart contracts reduce redundant operations by logistics providers, lowering the risk premium of financial institutions. Financial institutions can obtain real-time sales data from e-commerce platforms and inventory information from logistics providers through blockchain, accurately assess enterprise credit, and reduce default risks. Blockchain smart contracts can automatically trigger financing disbursements and repayment processes, reducing the manual intervention costs of financial institutions.

Proposition 6: Under cooperative game, $p_e^{BC1^*}, p_{de}^{BC2^*}, p_{sde}^{BC3^*}$ decreases as η increases, and $\tau^{BC1^*}, \tau^{BC2^*}, \tau^{BC3^*}$ is also negatively correlated with η.

E-commerce supply chain enterprises, under risk-averse conditions, tend to adopt conservative strategies. To better maintain their profit levels, e-commerce platforms are inclined to set lower product selling prices to attract more customers, obtain larger cash flows, and enhance their ability to cope with uncertain risks. Under the cooperation model between suppliers and logistics providers, suppliers may transfer the conservative pricing pressure of retailers by lowering wholesale prices, ultimately leading to a decrease in retail prices. In the alliance model between logistics providers and e-commerce platforms, the economies of scale of e-commerce platforms enable them to quickly respond to risk-averse behaviors through dynamic pricing models, thereby lowering retail prices. In the model of a three-party alliance, the reduced intermediate costs between e-commerce supply chain enterprises lead to lower sales prices. Highly risk-averse retailers may use blockchain technology to enhance the credibility of product quality, but consumers' willingness to pay for "safety premiums" is limited, and the final retail prices still tend to decline. The adoption of blockchain will increase the transparency of information, providing financial institutions with a clearer understanding of the market. Under risk-averse conditions, enterprises will set lower interest

rates to better alleviate financing risks. Financial institutions attract suppliers with higher credit ratings by lowering interest rates, reducing the probability of bad loans. Conservative financial institutions focus more on short-term liquidity and tend to provide short-term, low-interest loans. The high risk-averse behavior of suppliers and logistics providers may force financial institutions to lower financing interest rates to maintain cooperation, but this also compresses the profit margins of financial institutions. The data transparency of e-commerce platforms reduces information asymmetry, allowing financial institutions to offer lower interest rates based on transaction data recorded on the blockchain. During periods of high economic uncertainty, financial institutions may stabilize market confidence through interest rate cuts, further reinforcing the negative correlation between interest rates and the risk aversion coefficient.

TABLE 2 Shapley value of suppliers.

S	$\{s\}$	$\{s,d\}$	$\{s,e\}$	$\{s,d,e\}$
$v(S)$	$\overline{\pi}_s^{B^*}$	$\overline{\pi}_{sd}^{B^*}$	0	$\overline{\pi}_{se}^{B^*}$
$v(S\backslash s)$	0	$\overline{\pi}_d^{B^*}$	$\overline{\pi}_e^{B^*}$	$\overline{\pi}_{de}^{B^*}$
$v(S)-v(S\backslash s)$	$\overline{\pi}_s^{B^*}$	$\overline{\pi}_{sd}^{B^*}-\overline{\pi}_d^{B^*}$	$-\overline{\pi}_e^{B^*}$	$\overline{\pi}_{se}^{B^*}-\overline{\pi}_{de}^{B^*}$
$\lvert S\rvert$	1	2	2	3
$w(\lvert S\rvert)$	1/3	1/6	1/6	1/3
$w(\lvert S\rvert)\{v(S)-v(S\backslash s)\}$	$\frac{1}{3}\overline{\pi}_s^{B^*}$	$\frac{1}{6}\left(\overline{\pi}_{sd}^{B^*}-\overline{\pi}_d^{B^*}\right)$	$-\frac{1}{6}\overline{\pi}_e^{B^*}$	$\frac{1}{3}\left(\overline{\pi}_{se}^{B^*}-\overline{\pi}_{de}^{B^*}\right)$

6 Financing coordination strategy under the cooperative game after adopting blockchain

Through the above analysis, in the non-cooperative game without participating in the alliance, each member aims at maximizing their own interests, which will damage all gain levels. Some members will get higher profits, but the profit level of other members will be damaged. Through the alliance, we can integrate resources, lower the overall price and increase sales, thus increasing all gain level. Therefore, in the decision-making process, it is necessary to join a coordination mechanism to encourage members to cooperate and make joint decisions to achieve Pareto to pity. To promote e-commerce supply chain members to achieve top decision-making, it is the key to setting up reasonable coordination strategies.

6.1 Revenue sharing contract

The Shapley value, introduced by scholar Shapley in 1953, is employed to coordinate financing benefits. In cooperative game theory, the Shapley value method enables equitable profit distribution according to each member's contribution, thereby achieving supply chain coordination. Following Shapley [43], the Shapley value is calculated in Equation 33:

$$x_i = \sum_{i\in S, S\in N} w(|S|)[v(S) - v(S\backslash i)] \tag{33}$$

Where, i represents individual participants (suppliers, logistics providers, e-commerce platforms), $|S|$ denotes the number of participants in coalition S, $w(|S|)$ is the weighting factor, $v(S)$ represents the coalition's profit with participant i, $v(S\backslash i)$ indicates the coalition's profit without participant i, $v(S) - v(S\backslash i)$ measures the marginal contribution.

Under the cooperative game, the income function satisfies the convex assumption, which can effectively ensure the stability of income distribution. S Satisfy $\sum_{i\in S} N_i \geq v(S)$ and $\sum_{i=1}^{3} N_i = v(1,2,3)$ for any subset. Shapley value can ensure that each member's income in the alliance is not lower than that when he acts independently. That is, it satisfies super additivity, $v(S\cup T) \geq v(S) + v(T)$. Due to the convexity of Shapley value, the income distribution falls within the core, and there is no subset that can get higher income through self-association, thus preventing members from defecting. Shapley value ensures the complete distribution of the total income of the alliance and avoids the distribution dispute caused by the residual income. By calculating the average marginal contribution of members in all possible sub-alliances, Shapley value ensures the fairness principle that the greater the contribution, the higher the income, and reduces the possibility of defection caused by unfairness.

Based on the calculation formula, the Shapley for the vendor can be calculated in Table 2.

Substituting the calculation formula, the Shapley value method yields the allocation coefficients for the supplier, logistics company, and e-commerce regarding all gain as (0.385, 0.448, 0.167). We varied each coalition's value within ±20% around the baseline values while maintaining core properties such as super additivity and efficiency. For each set of adjusted payoffs, we recalculated the Shapley values using the standard formula, ensuring that the sensitivity analysis aligns with the theoretical framework of cooperative game theory. We conducted 100 Monte Carlo simulations, perturbing the payoffs with random noise to assess the distribution of Shapley values. The results show that the allocation (0.386, 0.458, 0.156) falls within the 95% confidence interval of the simulated values. The sensitivity analysis demonstrates that the Shapley allocation is not arbitrary but robust to reasonable fluctuations in the coalition's payoffs. This aligns with the economic principle that stable allocations require minimal deviations from baseline assumptions. The calculated Shapley values satisfy individual rationality and collective efficiency.

6.2 Cost sharing contract

To coordinate the relationships between e-commerce supply chain enterprises, referring to research of scholars such as Chu et al. [44], Ghosh et al. [45] and He et al. [46], a cost-sharing contract has been introduced. To promote cooperation among e-commerce supply chain enterprises, logistics providers propose corresponding financing solutions. Logistics providers are willing to share the financing costs of suppliers to attract them to choose themselves as their long-term partners responsible for transporting the products

produced by the suppliers. The proportion of financing costs that the logistics provider bears for the supplier is θ.

At this point, the financing profit functions for the supplier, the logistics provider, e-commerce and banks are shown in Equations 34–37:

$$\max_{w_s^B} \pi_s^{BS} = \left[w_s^{BS} - c_s^{BS} - (1-\theta)c_s^{BS}\tau^{BS} - b^{BS}\right]D^{BS} \tag{34}$$

$$\max_{w_d^B,\theta} \pi_d^{BS} = \left(w_d^{BS} - c_d^{BS} + \theta c_s^{BS}\tau^{BS} - b^{BS}\right)D^{BS} \tag{35}$$

$$\max_{p^N} \pi_e^{BS} = \left(p_e^{BS} - w_d^{BS} - w_s^{BS} - b^{BS}\right)D^{BS} \tag{36}$$

$$\max_{\tau^{BS}} \pi_b^{BS} = c_s^{BS}\tau^{BS}D^{BS} \tag{37}$$

where $\overline{D}^B = n\left(1 - p^B + a\xi + \epsilon\right)$.

According to the financing profit function, the expected profits and variances are respectively: $E\left(\pi_s^{BS}\right) = \left[w_s^{BS} - c_s^{BS} - (1-\theta)c_s^{BS}\tau^{BS} - b^{BS}\right]n\left(1 - p^{BS} + a\xi\right)$, $E\left(\pi_d^{BS}\right) = \left(w_d^{BS} - c_d^{BS} + \theta c_s^{BS}\tau^{BS} - b^{BS}\right)n\left(1 - p^{BS} + a\xi\right)$, $E\left(\pi_e^{BS}\right) = \left(p_e^{BS} - w_d^{BS} - w_s^{BS} - b^{BS}\right)n\left(1 - p^{BS} + a\xi\right)$, $var\left(\pi_s^B\right) = \left[w_s^{BS} - c_s^{BS} - (1-\theta)c_s^{BS}\tau^{BS} - b^{BS}\right]^2$, $var\left(\pi_d^{BS}\right) = \left(w_d^{BS} - c_d^{BS} + \theta c_s^{BS}\tau^{BS} - b^{BS}\right)^2$, $var\left(\pi_e^{BS}\right) = \left(p_e^{BS} - w_d^{BS} - w_s^{BS} - b^{BS}\right)^2$. Substituting these into the mean - standard deviation function, we can obtain: $U\left(\pi_s^{BS}\right) = \left[w_s^{BS} - c_s^{BS} - (1-\theta)c_s^{BS}\tau^{BS} - b^{BS}\right]n\left(1 - p^{BS} + a\xi - \epsilon\sigma\right)$, $U\left(\pi_d^{BS}\right) = \left(w_d^{BS} - c_d^{BS} + \theta c_s^{BS}\tau^{BS} - b^{BS}\right)n\left(1 - p^{BS} + a\xi - \epsilon\sigma\right)$, $U\left(\pi_e^{BS}\right) = \left(p_e^{BS} - w_d^{BS} - w_s^{BS} - b^{BS}\right)n\left(1 - p^{BS} + a\xi - \epsilon\sigma\right)$.

Proposition 7: According to Stackelberg, the best wholesale price of suppliers is $w_s^{BS^*} = \frac{c_s(\tau+1) - b - c_d + c_s\theta\tau + a\xi - \eta\sigma + 1}{2}$. The best logistics fee of logistics providers is $w_d^{BS^*} = \frac{c_d - c_s\theta\tau - w_s + a\xi - \eta\sigma + 1}{2}$, and the best cost sharing ratio is $\theta^{BS^*} = \frac{c_s\eta\tau(4b + c_d + w_s + c_d)}{2{c_s}^2\eta\tau^2}$. The best retail price of e-commerce platform is $p_e^{BS^*} = \frac{b + w_s + w_d + a\xi - \eta\sigma + 1}{2}$. The best financing interest rate of financial institutions is $\tau^{BS^*} = \frac{b + c_d - (a\xi - \eta\sigma + 1)}{2c_s\theta}$.

The top financing profits for each member are shown in Equations 38–41 respectively:

$$\pi_s^{BS*} = \frac{n\left(b + c_s - w_s + c_s\tau + c_s\theta\tau\right)\left(\begin{array}{l} 2b + c_d - \theta + w_s + 3b\theta + c_s\theta + c_d\theta - a\xi \\ +\eta\sigma - c_s\theta^2\tau - a\xi\theta - c_s\theta\tau + \sigma\theta\tau - 1 \end{array}\right)}{16\theta} \tag{38}$$

$$\pi_d^{BS*} = \frac{c_s\eta\theta\tau\left(b + c_s - w_s + c_s\tau\right)}{8} \tag{39}$$

$$\pi_e^{BS*} = \frac{n\left(2b + c_d + w_s - a\xi + \eta\sigma - c_s\theta\tau - 1\right)\left(2c_s - c_d - 3w_s + a\xi + 2c_s\tau - \eta\sigma + c_s\theta\tau + 1\right)}{2} \tag{40}$$

$$\pi_b^{BS*} = \frac{n\left(b + c_s - w_s + c_s\tau + c_s\theta\tau\right)\left(2b + c_d + w_s - a\xi + \eta\sigma - 1\right)}{16\theta} \tag{41}$$

7 Numerical analysis

The introduction of blockchain technology provides significant convenience for financing. For the guarantee financing model, blockchain can achieve the transmission of credit and information sharing, thereby improving financing efficiency. We explore the influencing factors of financing decisions, referring to the research of scholars such as Wu et al. [47], An et al. [48] and Qin et al. [49], setting $n = 2000, \tau = 0.03, \xi = 10, c_s^N = 0.3, c_d^N = 0.1, c_e^N = 0.5, c_s^B = 0.4, c_d^B = 0.1, c_e^B = 0.6$.

7.1 Influence of information verification

To explore the influence of information verification cost on the financing equilibrium strategy, sensitivity analysis is carried out, and the results are shown in Figure 2.

For suppliers, financing returns decrease with the rise in information verification costs, but the wholesale price raises with the rise in information verification costs. Higher information verification costs are detrimental to the improvement of suppliers' financing returns because they increase the intermediate costs of suppliers, leading to a reduction in financing returns. With higher information verification costs, suppliers will also set higher wholesale prices, making the price at which they sell to downstream enterprises relatively higher. Managers should appropriately reduce information verification costs during operations, as lower information verification costs are more beneficial to the development of suppliers.

For logistics providers, information verification costs indirectly contribute to the increase in their financing returns but have little impact on the guaranteed interest rate. With higher information verification costs, transaction information and logistics information become more transparent, and logistics providers are more willing to provide guarantees for suppliers, thereby increasing their own financing returns. However, the guaranteed interest rate of logistics providers is not affected by the cost of information verification efforts; as information verification costs increase, the guaranteed interest rate remains unchanged. Managers should reasonably formulate financing decisions in practice.

For e-commerce platforms, financing returns decrease with the increase in information verification effort costs, but retail prices are not affected by information verification costs. Higher information verification efforts increase the intermediate costs of e-commerce platforms, thereby affecting the improvement of profit levels. However, retail prices are not influenced by the information verification effort costs incurred by suppliers during the financing process. Managers should focus on actively leveraging technological advantages to reduce information verification effort costs.

For financial institutions, financing returns decrease with the increase in information verification effort costs, but the financing interest rate is not affected by the information verification effort costs. Higher information verification effort costs lead to increased costs for enterprises during the finance process, weakening the enthusiasm for financing and resulting in reduced financing amounts, which bring limited financing returns to financial institutions. The financing interest rate is influenced by various market factors and is less affected by the degree of enterprise information verification efforts; it does not change with the increase in information verification costs.

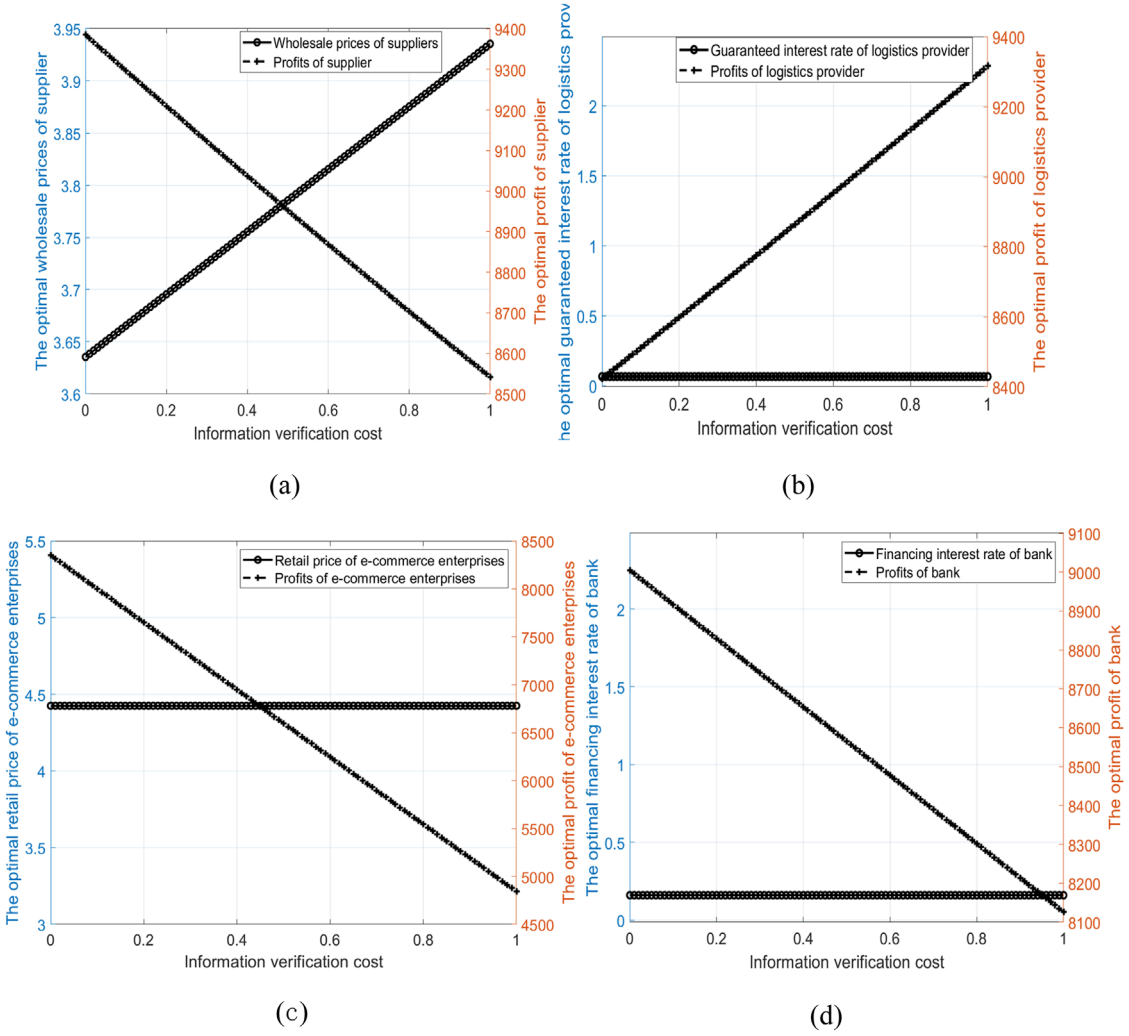


FIGURE 2
Influence of information verification. **(a)** Influence on suppliers. **(b)** Influence on logistics providers. **(c)** Influence on e-commerce enterprise. **(d)** Influence on bank.

7.2 Influence of default probability

The default probability of suppliers has significant implications for e-commerce supply chain financing decisions. A sensitivity analysis was conducted on this factor, and the results are shown in Figure 3.

For suppliers, financing returns decreases with the raise in the supplier's default probability, and a higher default probability also leads to a decrease in the supplier's wholesale price. A higher default probability is detrimental to the improvement of suppliers' financing returns and affects their reputation, leading to logistics providers being unwilling to provide guarantees for them, and thus failing to better alleviate their financial constraints. The wholesale price of suppliers also decreases with the raise in the default probability. Under a higher default probability, the cooperation between suppliers and e-commerce platforms is impaired, and suppliers can only adopt a price-cutting strategy to achieve better sales volumes and quickly recover funds.

For logistics providers, financing returns decrease with the rise in the default probability, but the guaranteed interest rate rises with the rise in the default probability. Under a higher default probability of suppliers, the financing returns of logistics providers are impaired. If logistics providers provide guarantees for suppliers and the suppliers fail to pay the principal and interest on time, the logistics providers need to cover the remaining payments. Under a higher default probability, the risks faced by logistics providers intensify, and the guaranteed interest rate will further increase to better protect their legitimate rights and interests.

For e-commerce platforms, financing returns paradoxically raise with the increase in the default probability, but retail prices are not affected by the default probability. Under higher default probability, suppliers set relatively lower wholesale prices, thereby reducing the procurement costs of e-commerce platforms and prompting an increase in returns. Regardless of the default probability, the sales prices of e-commerce platforms are not affected. Lower procurement costs will drive up the profit levels of e-commerce platforms.

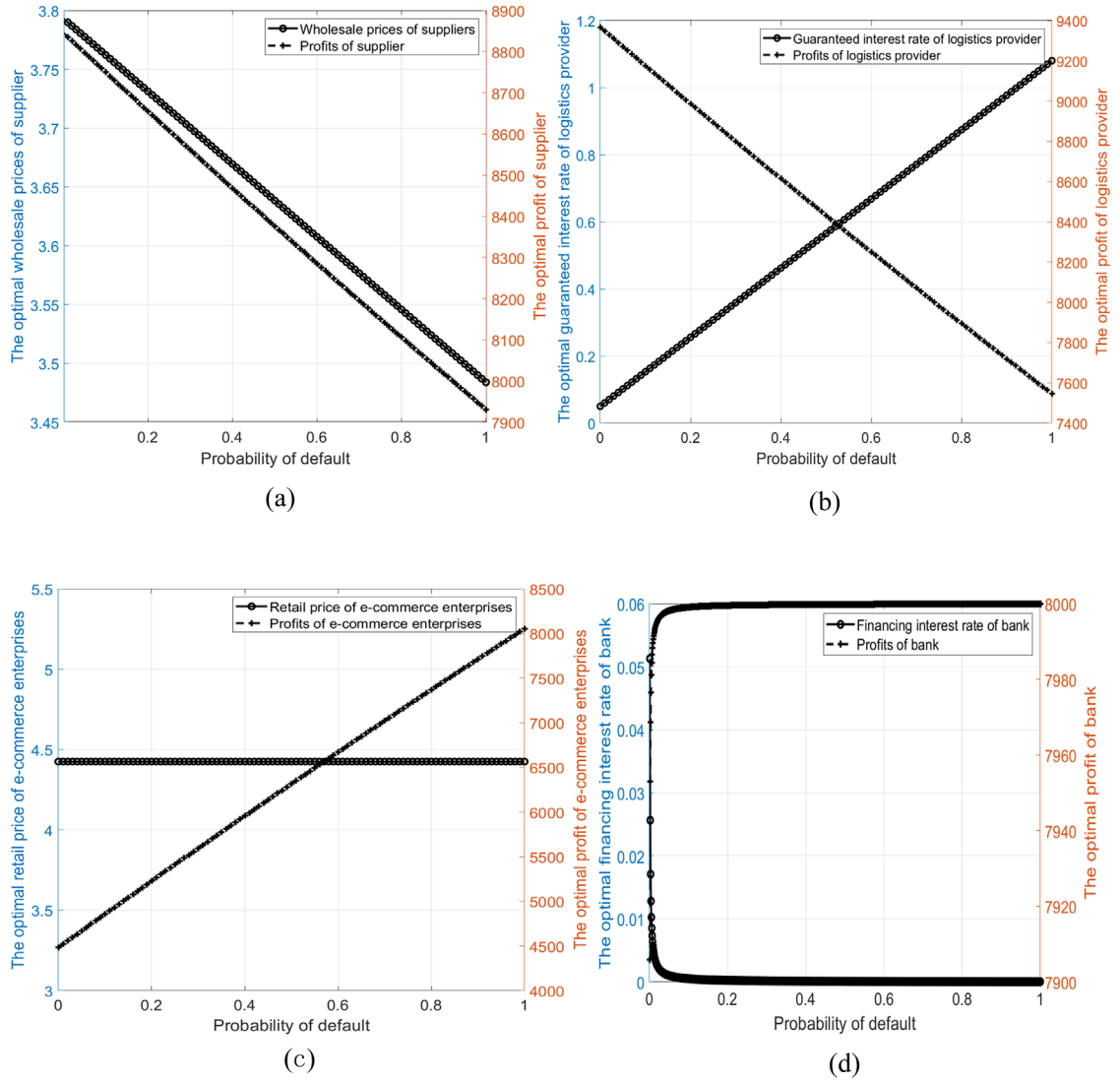


FIGURE 3
Influence of default probability. **(a)** Influence on suppliers. **(b)** Influence on logistics providers. **(c)** Influence on e-commerce enterprise. **(d)** Influence on bank.

For financial institutions, financing returns paradoxically raise with the rise in the default probability, but the finance interest rate decreases with the rise in the default probability. Under a higher default probability, after logistics providers provide guarantees for capital-constrained suppliers, they substitute in repaying the amounts not paid by suppliers, effectively safeguarding the rights and profit levels of financial institutions. To protect the interests of e-commerce supply chain enterprises, financial institutions also set relatively lower financing interest rates.

7.3 Influence of Blockchain's adoption

The cost of blockchain technology refers to the node deployment and maintenance costs under the consortium blockchain architecture. To study the influence of adopting blockchain on the finance returns of e-commerce supply chains, a sensitivity analysis is conducted. The effects of adopting blockchain are shown in Figure 4.

For suppliers, the adoption of blockchain increases their own costs, leading to higher wholesale prices. A lower cost of adopting blockchain technology helps improve the finance returns of suppliers. However, as the cost of blockchain increases, the profit of suppliers shows a trend of first decreasing and then increasing. At a higher level of blockchain adoption, blockchain effectively improves information transparency, increases trust among enterprises, simplifies the financing process, reduces intermediary costs, and enhances financing returns. When the blockchain cost is within a certain threshold range, the cost burden effect is significant, and returns gradually decrease. When the cost exceeds a certain threshold, the trust premium effect dominates, and the increase in trust effectively boosts financing returns.

For logistics providers, the introduction of blockchain technology prompts them to increase guarantee fees. Under

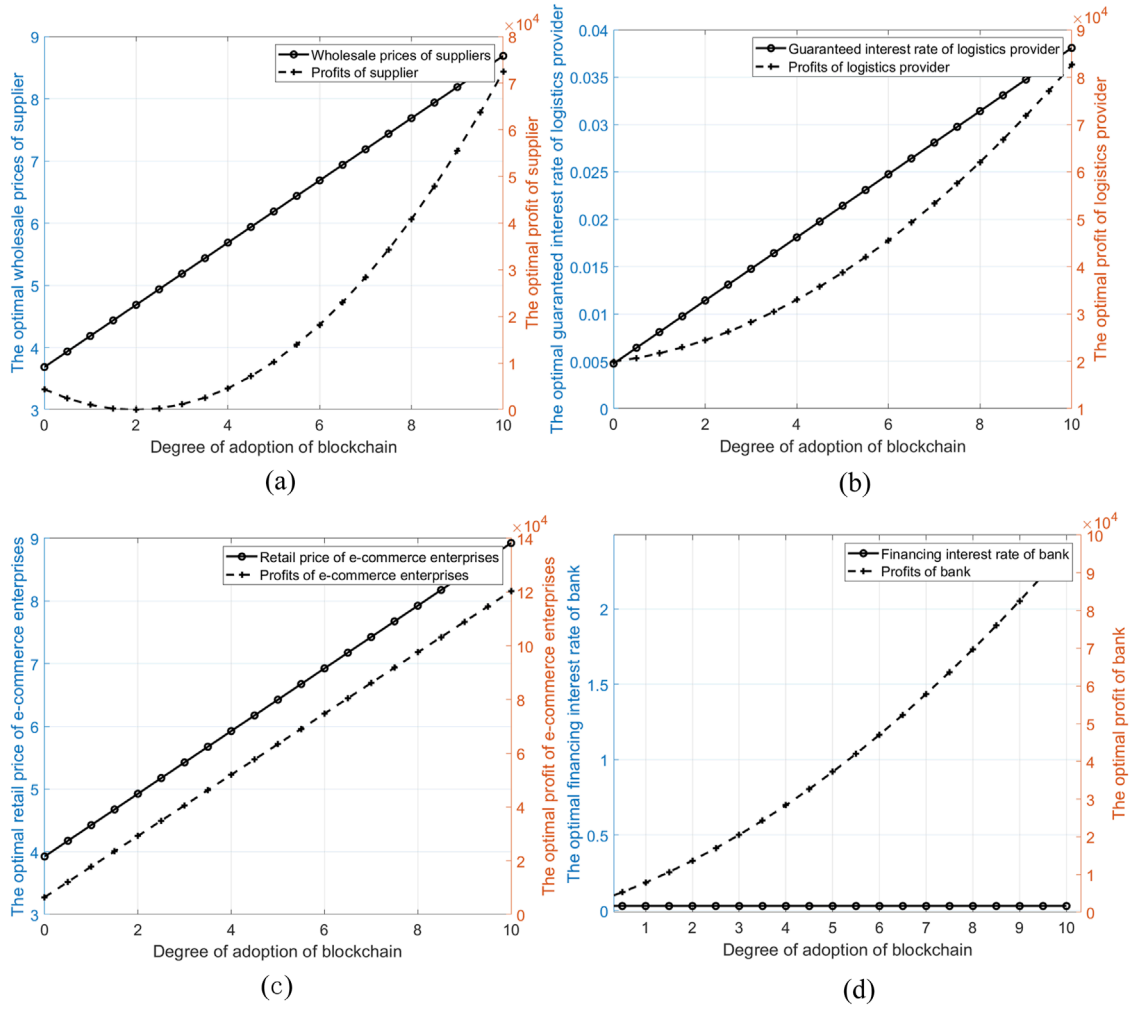


FIGURE 4
Influence of blockchain. **(a)** Influence on suppliers. **(b)** Influence on logistics providers. **(c)** Influence on e-commerce enterprise. **(d)** Influence on bank.

blockchain, logistics providers incur more costs when providing guarantees to suppliers, leading to higher guarantee fees. However, the financing returns of logistics providers increase with the degree of blockchain adoption, and the higher the degree of blockchain adoption, the greater the impact on the financing returns of logistics providers. Managers should actively introduce blockchain and raise the degree of adoption.

For e-commerce enterprises, the adoption of blockchain increases operating costs. Risk-averse e-commerce platforms typically consider adopting conservative operating strategies and therefore raising retail prices appropriately when costs increase, leading to higher retail prices for consumers. Blockchain promotes trust transmission and sharing along the supply chain, helping to attract more consumers and thus increase overall returns. Returns increase with the increased degree of blockchain technology adoption. Managers should actively adopt blockchain technology, increase technological investment, and gain more returns.

For financial institutions, the finance interest rate does not change with the degree of blockchain adoption; the rate remains constant. However, overall returns rise with the increased degree of blockchain adoption because, under the influence of blockchain, the market demand of suppliers increases correspondingly, the demand for funds gradually grows, loan demand intensifies, and thus the financing returns of financial institutions increase. Financial institutions should actively introduce blockchain technology and use technological advantages to boost returns.

7.4 Influence of risk aversion coefficient

To investigate how risk aversion before the adoption of blockchain affects financing decisions, a sensitivity analysis was conducted, and the results are shown in Figure 5.

For suppliers, before the adoption of blockchain, financing returns decrease with the rise in the risk aversion coefficient, while the wholesale price rises with the rise in the risk aversion coefficient. Under a higher risk aversion coefficient, suppliers, to protect their own interests, will set relatively higher wholesale prices to better cope with uncertain risks. However, higher sales prices will also

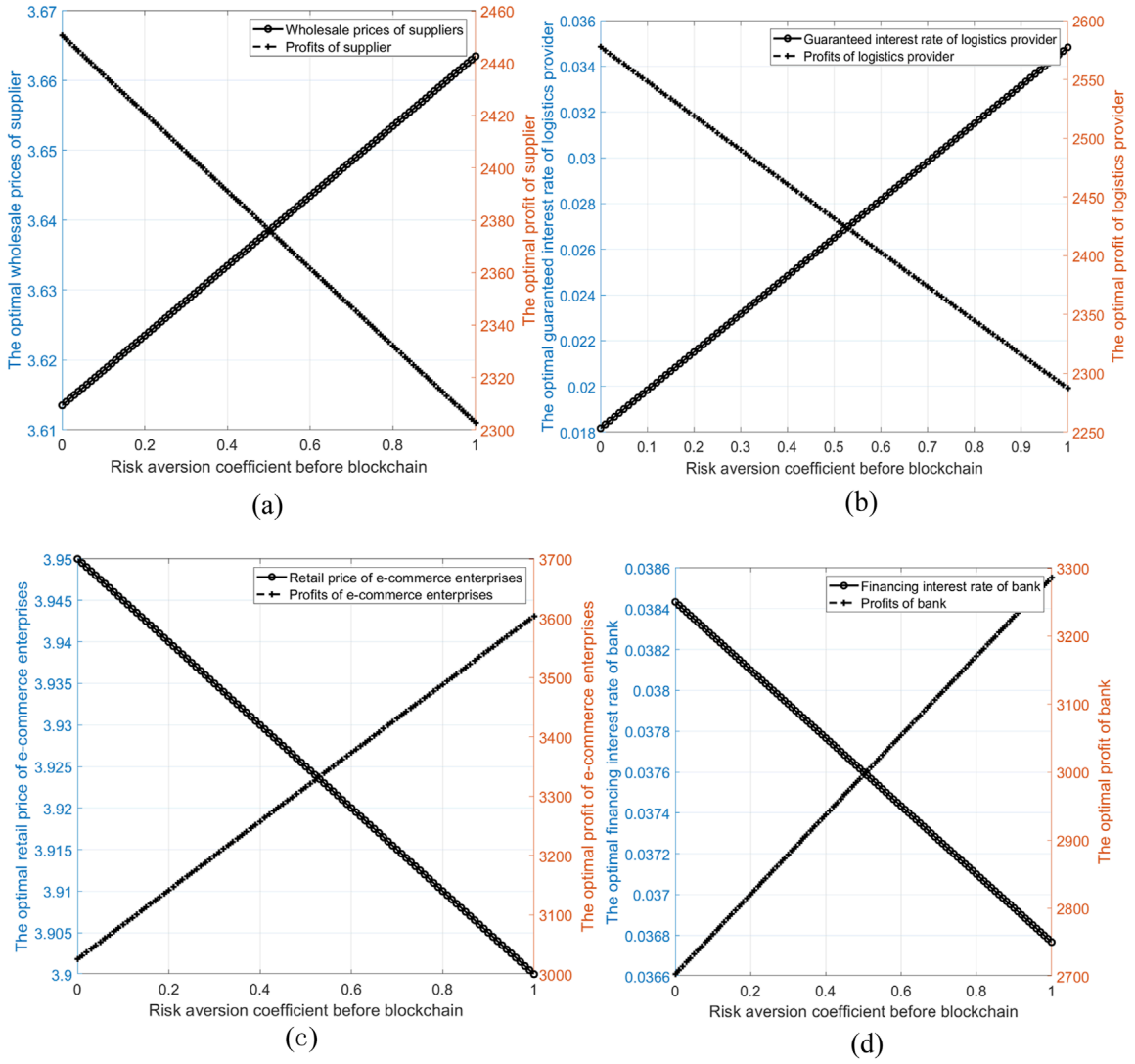


FIGURE 5
Influence of the risk aversion coefficient before blockchain adoption. **(a)** Influence on suppliers. **(b)** Influence on logistics providers. **(c)** Influence on e-commerce enterprise. **(d)** Influence on bank.

weaken consumers' purchasing enthusiasm, leading to a decrease in market demand, which in turn reduces the overall financing returns of suppliers.

For logistics providers, financing returns decrease with the rise in the risk aversion before the blockchain, while the guaranteed interest rate rises with the rise in the risk aversion coefficient. When logistics providers exhibit risk-averse behavior, they will set relatively higher guarantee interest rates to better avoid the default risk of suppliers and protect their own legitimate rights and interests. However, higher financing guarantee rates also lead to a reduction in the guaranteed amount requested by suppliers; risk-averse suppliers are unwilling to apply for more loans, resulting in a decrease in the overall financing returns of logistics providers.

For e-commerce platforms, returns raise with the rise in the risk aversion before blockchain adoption, while retail prices decrease with the rise in the risk aversion. When enterprises are risk-averse, they will appropriately lower sales prices to attract more buyers, which helps them to quickly recover funds. At the same time, this brings more revenue to e-commerce.

For financial institutions, financing returns rise with the rise in the risk aversion before blockchain adoption, while the finance interest rate decreases with the rise in the risk aversion coefficient. Under a higher risk aversion coefficient of enterprises and financial institutions, to better alleviate the financial constraints of suppliers, will moderately lower the finance interest rates, but the reduction is small because the financing interest rate is influenced by various factors. Meanwhile, suppliers are more willing to apply for financing, increase loan amounts, and financial institutions will obtain higher financing returns.

To analyze the influence of risk aversion, a analysis is conducted on the impact of risk aversion on the finance returns after blockchain. The results are shown in Figure 6.

For suppliers, when the degree of risk aversion is high, they tend to adopt more conservative sales strategies and appropriately lower

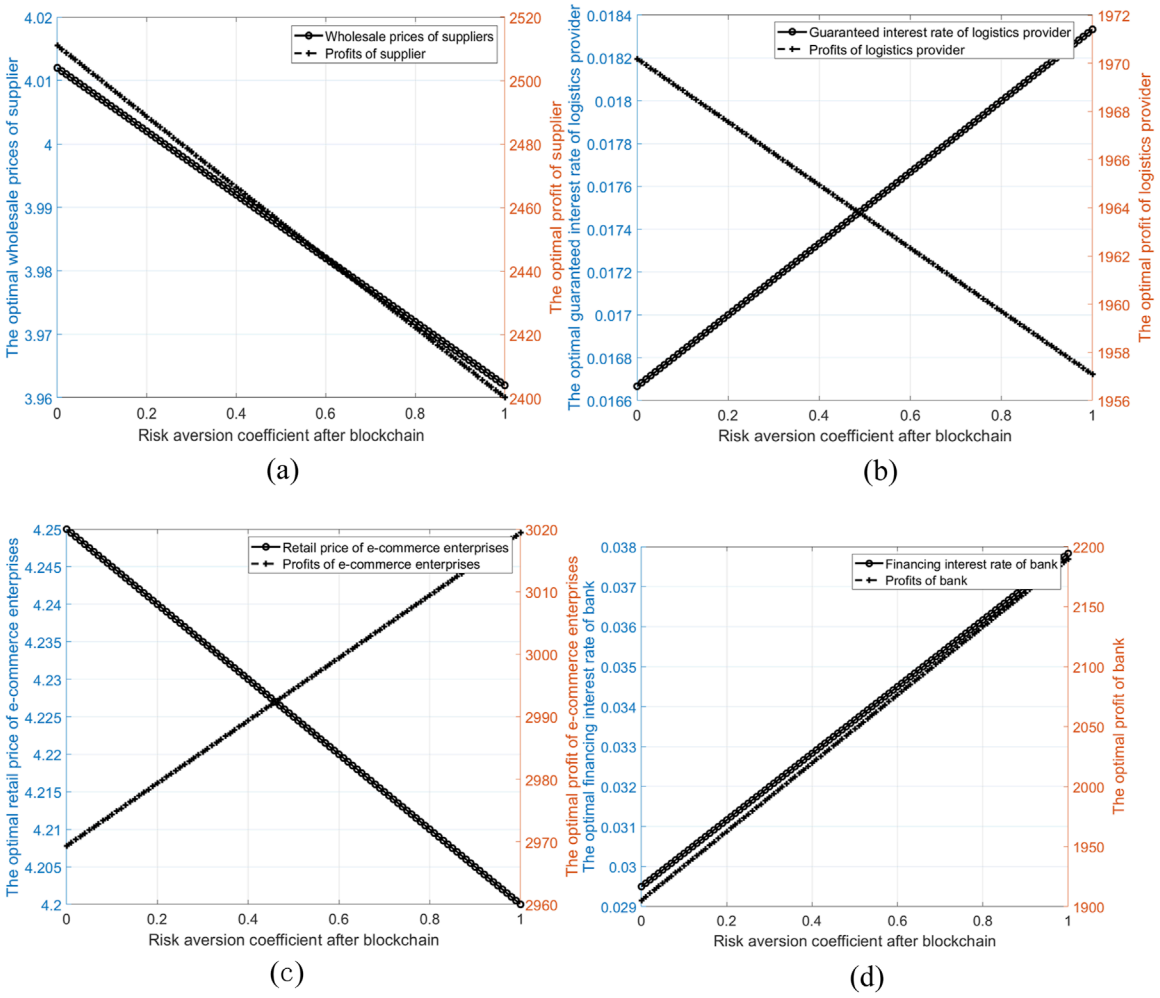


FIGURE 6
Influence of risk aversion. **(a)** Influence on suppliers. **(b)** Influence on logistics providers. **(c)** Influence on e-commerce enterprise. **(d)** Influence on bank.

wholesale prices to attract more consumers and maintain a certain sales volume. Blockchain enhances supply chain visibility and data sharing, thereby improving the accuracy of demand forecasting and prompting companies to adopt low-price sales strategies to quickly recover funds. A higher risk aversion does not necessarily help improve suppliers' finance returns; financing returns decrease with the rise in risk aversion. The higher the degree of risk aversion, the higher the guaranteed interest rates are set by logistics providers. Suppliers, to quickly recover funds, adopt low-price sales strategies, lower sales prices to attract more consumers, and expedite fund recovery, but this can, to some extent, impair the suppliers' returns. In practice, managers should moderately control the tendency towards risk aversion.

For logistics providers, when the degree of risk aversion is high, they will set higher guarantee interest rates to better protect their own legitimate rights and interests. However, higher guarantee interest rates can impair the enthusiasm of capital-constrained suppliers to apply for guarantees from logistics providers, reduce the number of guarantees applied for, and thus lead to a decrease in the financing returns of logistics providers. More risk aversion harms the returns of logistics providers; managers should strive to reduce their own risk aversion and set relatively lower acceptable guarantee interest rates to attract suppliers to apply for financing. Smart contracts can trigger automatic adjustments to production or procurement plans based on real-time demand signals. This reduces the possibility of excess inventory or stockouts, stabilizes product demand patterns, and better meets the company's risk aversion. The tamper-proof nature of blockchain establishes trust between supply chain partners. When companies trust the accuracy of shared data, they are more likely to adopt collaborative forecasting methods, better predict market demand, and further reduce risks.

For e-commerce platforms, the higher the risk aversion, the more inclined they are to set lower sales prices. By recording all transactions, logistics events, and inventory movements on an immutable ledger, blockchain achieves real-time visibility across the entire supply chain. This enables companies to access more granular historical and current data for more accurate demand forecasting. Under a higher risk aversion, suppliers' wholesale prices will be relatively lower, reducing the purchase costs for e-commerce platforms, and consequently, the retail prices sold to consumers will

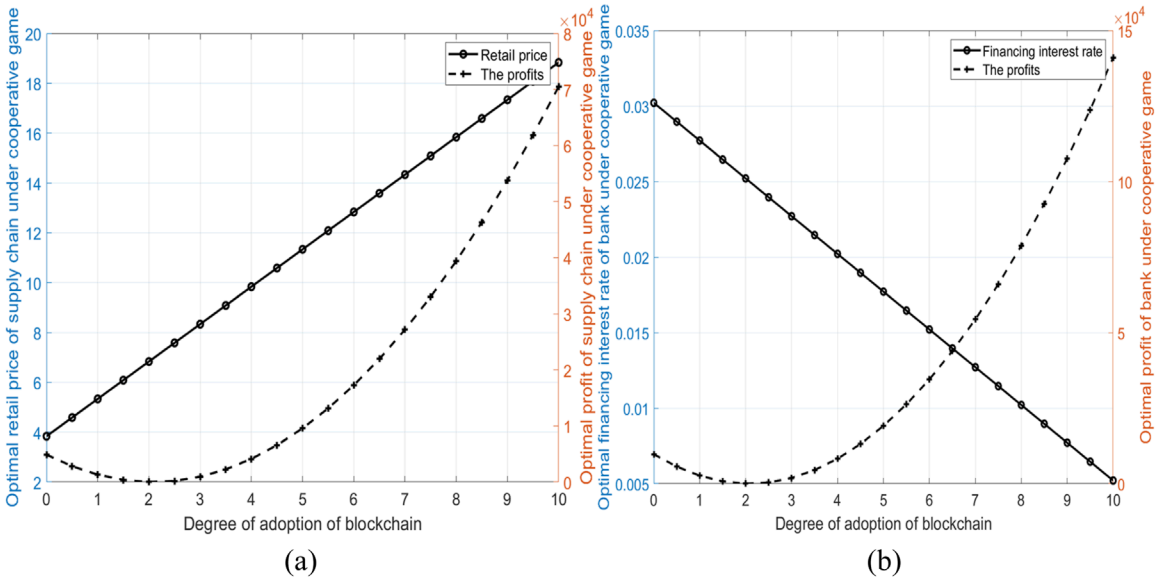


FIGURE 7
Influencing of blockchain under cooperative game. **(a)** Influence on supply chain. **(b)** Influence on bank.

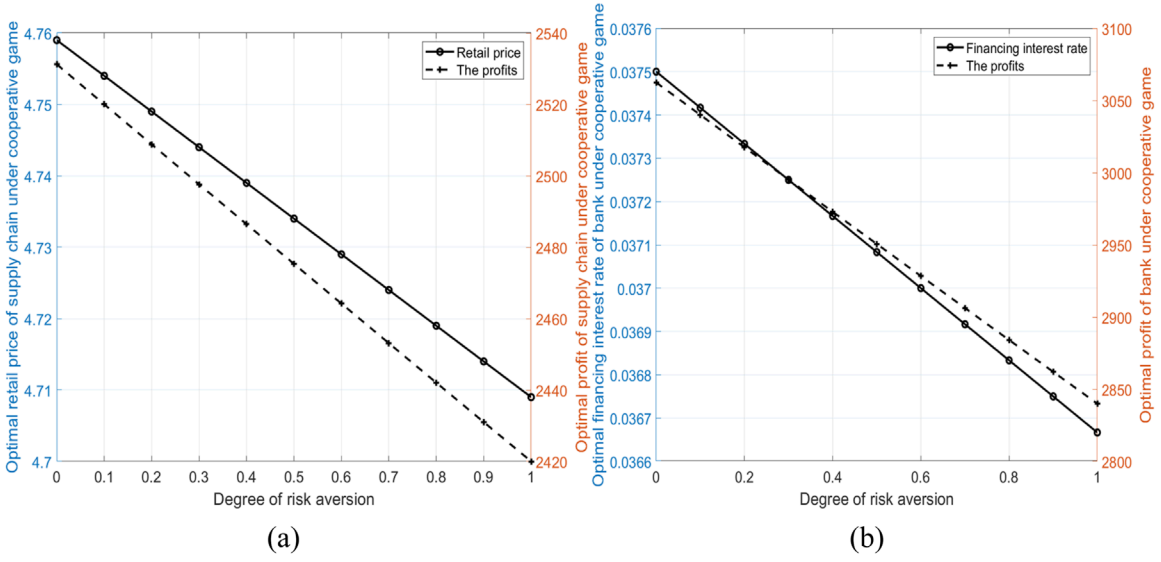


FIGURE 8
Influence of risk aversion under cooperative game. **(a)** Influence on supply chain. **(b)** Influence on bank.

also decrease. Lower sales prices can attract more consumers to buy products, leading to higher financing returns for e-commerce platforms. While maintaining risk aversion in daily operations, managers can set lower sales prices to attract more consumers and increase market demand.

For financial institutions, under a higher degree of risk aversion of enterprises, financial institutions will set relatively higher interest financing rates, but the variation in interest rates is not significant. When the risk aversion of enterprises is high, suppliers and logistics providers will adopt relatively conservative strategies, leading to lower financing returns and reduced cash flow. To maintain their own returns and reduce the risks faced, financial institutions will increase guarantee interest rates. Under higher loan interest rates, the returns of financial institutions will also be higher. Traditional supply chains suffer from fragmented data, leading to biased or incomplete predictions. The decentralized ledger of blockchain ensures that all stakeholders share consistent data access, thereby reducing prediction variance, increasing transparency, and making financial institutions more willing to provide funds to suppliers.

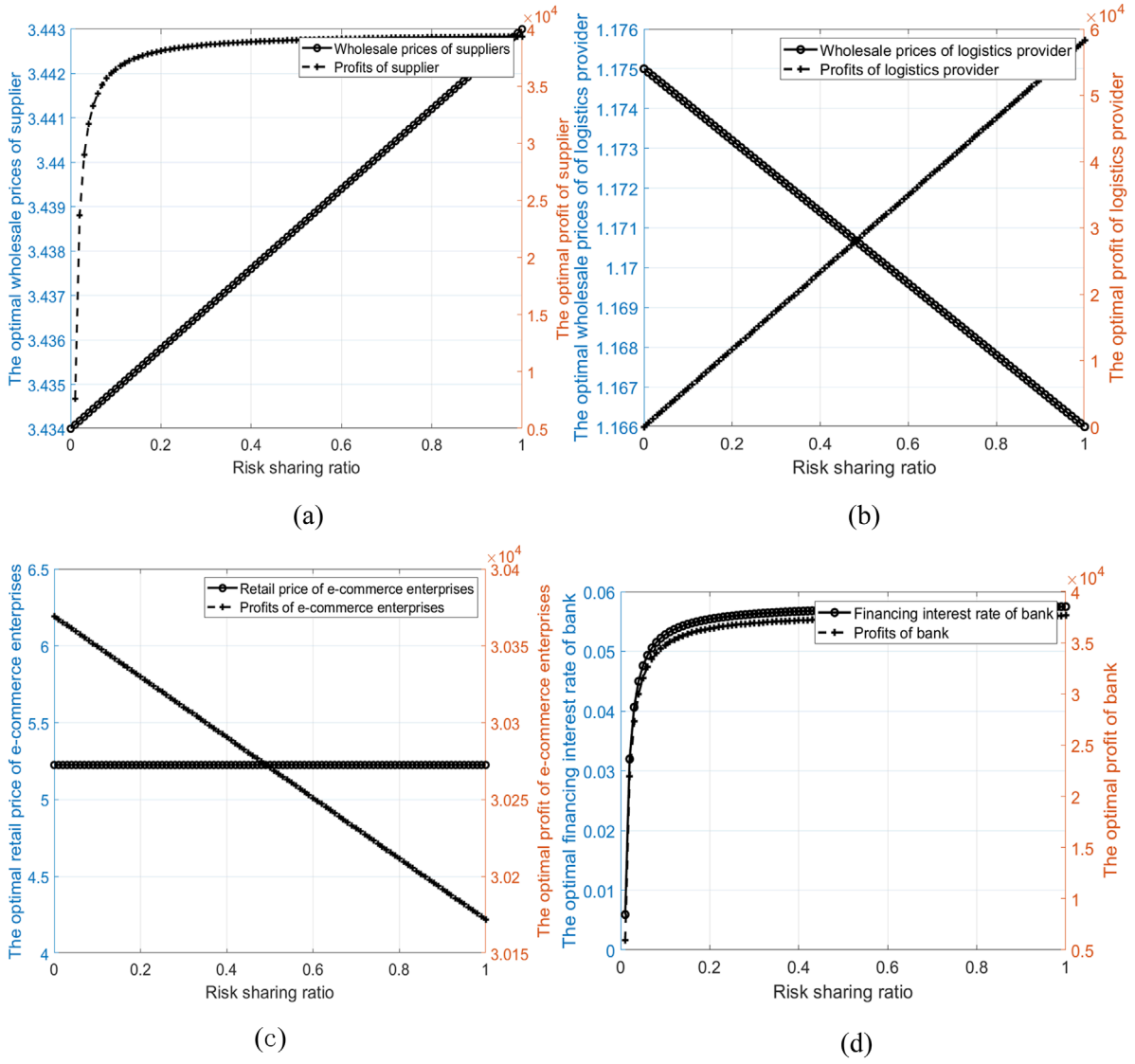


FIGURE 9
Influence of cost sharing. **(a)** Influence on suppliers. **(b)** Influence on logistics providers. **(c)** Influence on e-commerce enterprise. **(d)** Influence on bank.

7.5 Influence of blockchain and risk aversion under cooperative game

Under cooperative game theory, cooperation among suppliers, logistics providers, and e-commerce platforms is more helpful in increasing the profit.

First, we analyze the influence of blockchain on members under cooperative game theory. The results are shown in Figure 7.

For supply chain enterprises, under the three-party cooperation model, retail prices increase with the increased degree of blockchain adoption. The blockchain increases the financing costs and operating costs of enterprises, leading to relatively higher retail prices sold to consumers. The financing returns show a trend of first drop and then rise with the introduction of blockchain. Although blockchain technology increases costs in the short term, it also improves the traceability of information. Consumers trust enterprises more and are more willing to purchase products on the supply chain, thereby increasing the returns.

For financial institutions, under the cooperation model, financing interest rates decrease with the increased degree of blockchain adoption. Under higher blockchain, information transparency improves, allowing financial institutions to better review the qualifications of suppliers and increase their trust in suppliers. Therefore, they are more willing to provide financial support to suppliers. The financing returns of financial institutions show a trend of first drop and then rise with the increased degree of blockchain technology adoption. Under the blockchain background, financial institutions should increase the degree of blockchain adoption.

Next, we investigate the influence of risk aversion. The results are shown in Figure 8.

For e-commerce supply chain enterprises, under cooperative game theory, financing returns and sales prices both drop with the rise in risk aversion. Under a higher risk aversion, e-commerce supply chain enterprises set lower prices to attract more consumers and increase product sales. However, under a higher risk aversion coefficient, financing returns also drop with the rise in the risk aversion. Managers should reasonably control their own risk aversion behavior, as a higher risk aversion does not necessarily benefit the improvement of financing returns. It is also necessary to actively introduce blockchain to improve the transparency of the financing process and alleviate the risk avoidance of enterprises.

For financial institutions, in cooperative games, the greater the degree of risk aversion of enterprises, the lower the financing interest rate they will set, and at the same time, it will also lead to a decline in financing returns. When the risk aversion of an enterprise is relatively high, the enterprise will adopt a relatively conservative financing strategy. To attract suppliers with capital constraints, financial institutions will set a lower interest financing rate. The immutability of blockchain helps to improve the trust level of financial institutions in enterprises, and financial institutions reduce interest rates to ease the financial constraints of suppliers.

7.6 Influence of cost sharing coefficient

In supply chain coordination strategies, the risk-sharing coefficient affects the finance returns. The sensitivity analysis results are shown in Figure 9.

For suppliers, the bigger the risk-sharing coefficient, the more inclined suppliers are to set higher sales prices. Meanwhile, financing returns also rise with the rise in risk-sharing. Within a lower range of risk-sharing coefficients, the variation in financing returns is larger. As risk-sharing increases, the variation gradually decreases. In practice, managers should actively introduce risk-sharing mechanisms to alleviate their own financing constraints and improve financing returns.

For logistics providers, a certain degree of risk-sharing coefficient prompts logistics providers to reduce logistics costs and pass benefits to other enterprises. Under a higher risk-sharing coefficient, the financing returns of logistics providers also increase. Risk-sharing contracts help coordinate conflicts, better promote development, and improve financing returns. Certain incentive measures should be adopted to encourage managers to sign cost-sharing contract agreements and achieve the optimal results.

For e-commerce enterprises, the risk-sharing coefficient has no impact on retail prices. However, it leads to a decrease in returns. Under a certain risk-sharing mechanism, suppliers set higher sales prices, leading to reduced market consumer demand. The retail prices sold to consumers by e-commerce platforms do not increase, resulting in a decrease in their own returns.

For financial institutions, the financing interest rate will rise with the increase of risk sharing, and at the same time, the financing income will also raise with the raise of risk sharing. Under a higher cost-sharing ratio, financial institutions will set higher financing interest rates to increase their own financing returns.

8 Conclusion

Guarantee financing in SCF can effectively solve the funding issues of supply chain members by providing guarantees and applying for financing from financial institutions, thereby easing the capital constraints of vendors. However, as it develops, it gradually faces some bottleneck issues, such as information asymmetry among supply chain members, difficulties in credit transmission, and intensified fraud risks, all of which seriously hinder the development of guarantee financing. With the advancement of blockchain, its implementation in guarantee financing is comprehensive. Blockchain can help enterprises break through these bottlenecks and rais the efficiency and quality of financing.

This article takes guarantee financing in SCF as the object and analyzes the finance equilibrium and coordination of guarantee financing under blockchain. Based on the exposition of research backgrounds and current research status, the article first analyzes non-cooperative games, comparing the N mode without blockchain and the B mode with blockchain to explore the bearing of blockchain on the finance decisions. It then discusses cooperative game modes, studying three models: cooperation between vendor and logistics companies, cooperation between logistics companies and e-commerce platforms, and cooperation among all three, analyzing the impact of different cooperation modes on financing decisions after the blockchain. To coordinate the conflicts, the article introduces coordination strategies, discussing revenue sharing and cost-sharing separately to explore the bearing of coordination on the supply chain.

Through numerical examples, research finds that the cost of information verification will damage the financing income of suppliers, but it will promote the financing income of logistics providers. Supplier default will also damage the financing income of suppliers and logistics providers. The impact of blockchain on finance profits exhibits a "U"-shaped variation. When the adoption of blockchain exceeds a certain threshold range, it can effectively increase the financing profits of suppliers and logistics providers. A higher risk aversion coefficient can impair the financing profits of suppliers and logistics providers but helps to increase the profits of e-commerce and financial institutions. Compared to non-cooperative modes, cooperative modes are more helpful in increasing the overall profits. Introducing revenue-sharing and risk-sharing contracts can better alleviate the financial constraints of suppliers.

Based on the models and numerical results of this paper, for managers of supply chains and financial institutions, when considering the introduction of blockchain technology, managers need to comprehensively calculate the actual costs of blockchain construction and operation and match these with the risk-bearing capacities of their own enterprises and partners. If the organizational size, transaction frequency, and supply chain complexity are low, excessively high blockchain costs may weaken the marginal benefits of financing; however, if the cooperative ecosystem tends to mature and high transparency of the transaction chain is required, moderate resource investment on the chain will bring trust premium effects, significantly improving financing efficiency and negotiation positions. Managers should reasonably control their own risk aversion behavior and strengthen cooperation among enterprises.

Managers need to establish reasonable profit distribution and cost-sharing mechanisms among suppliers, logistics providers, and e-commerce/financial institutions. From the Shapley value allocation presented in this paper, through negotiation, Shapley allocation ratios based on convex profit functions should be set to ensure that all parties receive allocation shares at the game equilibrium point that are no less than their gains in any sub-alliance, thus avoiding the dissolution of alliances driven by self-interest. Managers should also set up corresponding incentive mechanisms, actively introduce cost-sharing contracts, and improve financing income.

Although this paper comprehensively discusses the financing equilibrium strategies under blockchain from theoretical and numerical perspectives, there are still limitations. When studying the behavioral preferences of managers, this paper only considers risk-averse behavior; future research could introduce overconfidence and other irrational preferences, which may significantly influence pricing and financing behaviors. Secondly, in constructing cooperative games and contract designs, the paper assumes that the cooperative profit function satisfies convexity and that members have complete consensus on transaction terms. Real-world collaborations involve factors such as negotiation costs, information lags, and trust deficits, which can weaken the effectiveness of contract execution. Future research should consider introducing incomplete information games, multi-stage evolutionary games, or vertical principal-agent models to meticulously discuss information asymmetry and multi-party negotiation mechanisms in the cooperative process.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

LY: Writing – original draft, Methodology, Writing – review and editing. QH: Supervision, Writing – review and editing, Conceptualization, Project administration.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Liaoning Provincial Department of Education's basic scientific research project in 2023, "Research on the path of matching supply and demand of smart pension industry under government subsidy mechanism" (JYTMS20231228).

Acknowledgments

We acknowledge all those who contributed to this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Chang S, Li A, Wang X, Zhang J. Optimal combination of platform channel contract and guarantee financing strategy in e-commerce market. *Transportation Res E: Logistics Transportation Rev* (2023) 172:103094. doi:10.1016/j.tre.2023.103094

2. Chen L, Chan HK, Zhao X. Supply chain finance: latest research topics and research opportunities. *Int J Prod Econ* (2020) 229:107766. doi:10.1016/j.ijpe.2020.107766

3. Yu Z (2022). *Supply chain finance, core company performance and financing difficulties of SMEs*.

4. He H, Li S, Hu L, Duarte N, Manta O, Yue XG. Risk factor identification of sustainable guarantee network based on logistic regression algorithm. *Sustainability* (2019) 11(13):3525. doi:10.3390/su11133525

5. Zhang H, Chen Z. Design of risk guarantee platform for financial support of port logistics industry. *J Coastal Res* (2020) 103(SI):749–52. doi:10.2112/si103-154.1

6. Li S, Qu S. The three-level supply chain finance collaboration under blockchain: income sharing with shapley value cooperative game. *Sustainability* (2023) 15(6):5367. doi:10.3390/su15065367

7. Liu J, Yan L, Wang D. A hybrid blockchain model for trusted data of supply chain finance. *Wireless Personal Commun* (2022) 127:919–43. doi:10.1007/s11277-021-08451-x

8. Bhatia MS, Chaudhuri A, Kayikci Y, Treiblmaier H. Implementation of blockchain-enabled supply chain finance solutions in the agricultural commodity supply chain: a transaction cost economics perspective. *Prod Plann Control* (2024) 35(12):1353–67. doi:10.1080/09537287.2023.2180685

9. Deng L, Li Y, Wang S, Luo J. The impact of blockchain on optimal incentive contracts for online supply chain finance. *Environ Sci Pollut Res* (2023) 30(5):12466–94. doi:10.1007/s11356-022-22498-8

10. Wang X, Zhao J, Zhang H, Tang X. The impact of the scale of third-party logistics guaranteeing firms on bank credit willingness in supply chain finance: an ERP study. *Front Psychol* (2022) 13:853888. doi:10.3389/fpsyg.2022.853888

11. Li S, Chen X. The role of supply chain finance in third-party logistics industry: a case study from China. *Int J Logistics Res Appl* (2018) 22(2):154–71. doi:10.1080/13675567.2018.1502745

12. Zhang C, Fan LW, Tian YX. Optimal operational strategies of capital-constrained supply chain with logistics service and price dependent demand under 3PL financing service. *Soft Comput* (2020) 24(4):2793–806. doi:10.1007/s00500-019-04500-7

13. Song X, Wu X, Liu C, Yang H. The order financing risk evaluation under the financial supply chain model based on the logistic model. *J Mathematics* (2022) 2022(1):1854436. doi:10.1155/2022/1854436

14. Chen X, Cai GG. Joint logistics and financial services by a 3PL firm. *Eur J Oper Res* (2011) 214(3):579–87. doi:10.1016/j.ejor.2011.05.010

15. Yu X, Chen H, Xiang K, Wan Z. Supply chain financing mechanism with guarantee insurance. *Managerial Decis Econ* (2021) 42(2):308–18. doi:10.1002/mde.3236

16. Liu L, Zhang W, Wang L, Sun Y. Financing provision strategies for third-party logistics with consideration of supply chain resilience. *Int Trans Oper Res* (2024). doi:10.1111/itor.13602

17. Bi G, Shen F, Xu Y. Third-party logistics firm's technology investment and financing options in platform-based supply chain with 4PL service. *Naval Res Logistics (Nrl)* (2024) 71(6):763–82. doi:10.1002/nav.22177

18. Sun J, Yuan P, Hua L. Pricing and financing strategies of a dual-channel supply chain with a capital-constrained manufacturer. *Ann Operations Res* (2023) 329(1):1241–61. doi:10.1007/s10479-022-04602-w

19. Li L, Wang Z, Zhao X. Configurations of financing instruments for supply chain cost reduction: evidence from Chinese manufacturing companies. *Int J Operations Prod Management* (2022) 42(9):1384–406. doi:10.1108/ijopm-12-2021-0755

20. Lai Z, Lou G, Zhang T, Fan T. Financing and coordination strategies for a manufacturer with limited operating and green innovation capital: bank credit financing versus supplier green investment. *Ann Operations Res* (2023) 331(1):85–119. doi:10.1007/s10479-021-04098-w

21. Tsai CH. Supply chain financing scheme based on blockchain technology from a business application perspective. *Ann Operations Res* (2023) 320(1):441–72. doi:10.1007/s10479-022-05033-3

22. Gong Y, Zhang T, Dong P, Chen X, Shi Y. Innovation adoption of blockchain technology in supply chain finance. *Prod Plann Control* (2024) 35(9):992–1008. doi:10.1080/09537287.2022.2155564

23. Tan J, Wong WP, Tan CK, Jomthanachai S, Lim CP. Blockchain-based logistics 4.0: enhancing performance of logistics service providers. *Asia Pac J Marketing Logistics* (2024) 36(6):1442–63. doi:10.1108/apjml-07-2023-0650

24. Raj PVRP, Jauhar SK, Ramkumar M, Pratap S. Procurement, traceability and advance cash credit payment transactions in supply chain using blockchain smart contracts. *Comput Ind Eng* (2022) 167:108038. doi:10.1016/j.cie.2022.108038

25. Li M, Shao S, Ye Q, Xu G, Huang GQ. Blockchain-enabled logistics finance execution platform for capital-constrained E-commerce retail. *Robotics and Computer-Integrated Manufacturing* (2020) 65:101962. doi:10.1016/j.rcim.2020.101962

26. Xu L, Yang Y, Chu X. Research on the influence mechanism of block chain on the credit of transportation capacity supply chain finance. *Math Probl Eng* (2021) 2021(1):1–11. doi:10.1155/2021/1812096

27. Harish AR, Liu XL, Li M, Zhong RY, Huang GQ. Blockchain-enabled digital assets tokenization for cyber-physical traceability in E-commerce logistics financing. *Comput Industry* (2023) 150:103956. doi:10.1016/j.compind.2023.103956

28. Wang L, Luo XR, Lee F, Benitez J. Value creation in blockchain-driven supply chain finance. *Inf Management* (2022) 59(7):103510. doi:10.1016/j.im.2021.103510

29. Rao Q, Li W. Risk evaluation and forecast behavior analysis of supply chain financing based on blockchain. *Wireless Commun Mobile Comput* (2022) 2022(1):1–11. doi:10.1155/2022/7668474

30. Jiang R, Kang Y, Liu Y, Liang Z, Duan Y, Sun Y, et al. A trust transitivity model of small and medium-sized manufacturing enterprises under blockchain-based supply chain finance. *Int J Prod Econ* (2022) 247:108469. doi:10.1016/j.ijpe.2022.108469

31. Ning L, Yuan Y. How blockchain impacts the supply chain finance platform business model reconfiguration. *Int J Logistics Res Appl* (2023) 26(9):1081–101. doi:10.1080/13675567.2021.2017419

32. Wang L, Wang Y. Supply chain financial service management system based on block chain IoT data sharing and edge computing. *Alexandria Eng J* (2022) 61(1):147–58. doi:10.1016/j.aej.2021.04.079

33. Liu L, Li Y, Jiang T. Optimal strategies for financing a three-level supply chain through blockchain platform finance. *Int J Prod Res* (2023) 61(11):3564–81. doi:10.1080/00207543.2021.2001601

34. Dong L, Qiu Y, Xu F. Blockchain-enabled deep-tier supply chain finance. *Manufacturing Serv Operations Management* (2023) 25(6):2021–37. doi:10.1287/msom.2022.1123

35. Yu Y, Huang G, Guo X. Financing strategy analysis for a multi-sided platform with blockchain technology. *Int J Prod Res* (2021) 59(15):4513–32. doi:10.1080/00207543.2020.1766718

36. Kucukaltan B, Kamasak R, Yalcinkaya B, Irani Z. Investigating the themes in supply chain finance: the emergence of blockchain as a disruptive technology. *Int J Prod Res* (2024) 62(22):8173–92. doi:10.1080/00207543.2022.2118886

37. Han S, Ulhøi JP, Song H. Digital trust in supply chain finance: the role of innovative fintech service provision. *J Enterprise Inf Management* (2024) 37(6):1737–62. doi:10.1108/JEIM-07-2022-0238

38. Pournader M, Shi Y, Seuring S, Koh SL. Blockchain applications in supply chains, transport and logistics: a systematic review of the literature. *Int J Prod Res* (2020) 58(7):2063–81. doi:10.1080/00207543.2019.1650976

39. Tian L, Vakharia AJ, Tan Y, Xu Y. Marketplace, reseller, or hybrid: strategic analysis of an emerging e-commerce model. *Prod Operations Manage* (2018) 27(8):1595–610. doi:10.1111/poms.12885

40. Li G, Tian L, Zheng H. Information sharing in an online marketplace with co-opetitive sellers. *Prod Operations Manage* (2021) 30(10):3713–34. doi:10.1111/poms.13460

41. Choi TM, Ma C, Shen B, Qi S. Optimal pricing in mass customization supply chains with risk-averse agents and retail competition. *Omega* (2019) 88:150–61. doi:10.1016/j.omega.2018.08.004

42. Gupta V, Ivanov D. Dual sourcing under supply disruption with risk-averse suppliers in the sharing economy. *Int J Prod Res* (2020) 58(1):291–307. doi:10.1080/00207543.2019.1686189

43. Shapley LS. *A value for n-person games*. Santa Monica, Calif: Rand Corp (1952). p. 13.

44. Chu LY, Sappington DEM. Simple cost-sharing contracts. *Am Econ Rev* (2007) 97(1):419–28. doi:10.1257/aer.97.1.419

45. Ghosh D, Shah J. Supply chain analysis under green sensitive consumer demand and cost sharing contract. *Int J Prod Econ* (2015) 164:319–29. doi:10.1016/j.ijpe.2014.11.005

46. He P, He Y, Shi CV, Xu H, Zhou L. Cost-sharing contract design in a low-carbon service supply chain. *Comput Ind Eng* (2020) 139:106160. doi:10.1016/j.cie.2019.106160

47. Wu C, Xu C, Zhao Q, Zhu J. Research on financing strategy under the integration of green supply chain and blockchain technology. *Comput and Ind Eng* (2023) 184:109598. doi:10.1016/j.cie.2023.109598

48. An S, Li B, Wang M, Zheng W. Financing strategy of the capital-constrained supply chain under uncertainty: the impact of blockchain technology on the credit period. *Ind Management Data Syst* (2023) 123(10):2659–80. doi:10.1108/imds-10-2022-0657

49. Qin J, Fu H, Wang Z, Lyu X. Financing decisions in capital-constrained supply chains considering uncertain emission reduction outputs and blockchain technology applications. *Int J Prod Res* (2025) 1–22. doi:10.1080/00207543.2025.2483432

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 26 August 2025
DOI 10.3389/fphy.2025.1651460

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Xin Xu,
Nanjing University of Finance and Economics, China
Tiago Novaes Mathias,
Port and Airport Research Institute, Japan
Yun Jiang,
China Special Vehicle Research Institute, China

*CORRESPONDENCE
Jianfeng Guo,
guojf@casipm.ac.cn

RECEIVED 21 June 2025
ACCEPTED 23 July 2025
PUBLISHED 26 August 2025

CITATION
Zhang X, Guo J and Liu S (2025) Structural robustness of China's UAV enterprise network: a national and provincial analysis based on business homogeneity. *Front. Phys.* 13:1651460.
doi: 10.3389/fphy.2025.1651460

COPYRIGHT
© 2025 Zhang, Guo and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Structural robustness of China's UAV enterprise network: a national and provincial analysis based on business homogeneity

Xuemei Zhang[1], Jianfeng Guo[2,3]* and Siyao Liu[2]

[1]College of Humanities and Social Sciences, Nanjing University of Aeronautics and Astronautics, Nanjing, China, [2]Institutes of Science and Development, Chinese Academy of Sciences, Beijing, China, [3]School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing, China

Introduction: The rapid growth of the unmanned aerial vehicle (UAV) market has surged enterprises, exposing them to systemic risks from business homogeneity. This issue has spread not only within the same level of the industrial chain but also across different segments. Despite China's enterprises spanning the entire chain as the largest market, their business structure and regional differences remain largely unexplored.

Method: This study constructs a three-tier UAV enterprise network based on overlapping business scopes among firms, forming upstream, midstream, and downstream layers, and examines its national and provincial structural robustness through topological characteristics and network dismantling experiments.

Results: The results reveal that national- and provincial-level UAV enterprise networks follow a power-law distribution, reflecting a "rich-get-richer" pattern. The national network is less robust than the provincial average, with downstream segments in delivery, agriculture, and other applications particularly vulnerable. Provincial networks exhibit strong regional heterogeneity; economically developed provinces demonstrate stronger internal coordination but are more fragile due to high centralization. Furthermore, provincial networks fall into four structural types with differing levels of efficiency and robustness.

Discussion: The study highlights how business similarity fosters local coordination but increases systemic risk through structural homogeneity. To enhance network resilience, especially in downstream and centralized regions, strategies such as modular design and region-specific coordination are essential.

KEYWORDS

unmanned aerial vehicle, multi-layer complex network, targeted attacks, topological characteristics, structural robustness

1 Introduction

In recent years, the global unmanned aerial vehicle (UAV), commonly known as the drone industry, has expanded rapidly, reaching $31.66 billion in 2023 and projected

to grow at a compound annual growth rate (CAGR) of 18.2% in 2024 [1]. The rapid growth has spurred in UAV-related enterprises [2, 3], many of which pursue economies of scale through shared resources, labor pools, and knowledge spillovers [4–6]. This has resulted in industrial clusters with high business homogeneity [7], which draws from institutional isomorphism theory. According to this theory, organizations operating within the same domain tend to become increasingly similar over time due to coercive, mimetic, and normative pressures, resulting in convergence in structure, strategy, and service offerings [8]. While such homogeneity may support upgrading in traditional industries [9, 10], it can pose systemic risk under external shocks in high-tech industries such as UAVs [10, 11]. Given China's dominant position in the global UAV market [12, 13], similar vulnerabilities are expected to emerge. However, the network structure and risk resilience of China' UAV enterprise network at both national and provincial levels remain unclear, as existing literature has yet to examine these aspects in detail.

Moreover, the risks brought by business homogeneity are not limited to a single stage of the industry chain but span across multiple levels [11, 14]. The UAV industry, like many manufacturing sectors, can be divided into upstream, midstream, and downstream segments, referring to suppliers, manufacturers, and distributors, respectively [15, 16]. The convergence of suppliers toward single-source materials introduces systemic risk [17]. When environmental or geopolitical disruptions exceed a certain threshold, failures can rapidly cascade from manufacturers to end users through downstream applications [17, 18]. Such cascade vulnerability was evident in the 2011 Japan earthquake, where the disruption of a single key supplier halted Toyota's production nationwide and impacted global end users [18]. Therefore, it is essential to systematically assess the structural robustness of China's UAV market to mitigate risks from a multi-level industry chain perspective.

The structural robustness, a concept originally developed in the field of structural engineering [19–21], has been widely adopted in network analysis [22–24]. Architectural literature defines robustness as a structure's insensitivity to local failure [19, 21], which is determined by both internal deterioration and external shocks [21, 25, 26]. Such vulnerability is often influenced by material properties and spatial arrangement [21, 26]. Similarly, in the enterprise network, structural robustness is shaped by business scopes and the positions of key business clusters. Previous research has shown that enterprises with similar business activities are more likely to cluster together [7, 27, 28]. Despite the potential significance of this relationship, existing research in both network structure and industrial value chain literature has yet to explore how business similarity influences structural robustness.

Methodologically, we employ a multi-layer complex network approach, which is increasingly applied to analyze trade flows [29–32]. In this study, we are the first to construct a three-layer complex network for UAV enterprise operations, segmented into upstream, midstream, and downstream tiers, a division consistent with the most widely recognized industrial chain structure. All 2,120 listed UAV-related enterprises comes from Qichacha, a professional business database that provides comprehensive firm-level information. Based on their registered business scope, enterprises were assigned to different stages of the industry chain. Details are recorded in Section 2.1. Data sources. The analysis focuses on China, given its dominant position in the global UAV market. To examine regional differences, we selected six key provinces in China. The six provinces are Guangdong, Jiangsu, Beijing, Zhejiang, Shandong, and Henan, accounting for approximately 60% of the sampled enterprise. A key concept in our analysis is business homogeneity, which we define as the semantic similarity between the business scope. This operationalization reflects the degree to which firms engage in similar activities or serve related functions along the industry chain.

In principle, our work renders three major contributions. First, We pioneer highlight the examination of business homogeneity in industrial network. Unlike previous studies that typically focus on trade flows [29–32], our research investigates the relationship between enterprises' business scopes. It reveals the importance of industrial clusters composed of similar businesses in the robustness of network structure, thereby significantly extending the research perspective and the boundaries of network analysis. Second, we adopt a multi-scale spatial framework to assess the structural robustness of UAV network at national and provincial levels, examining both normal topology and response to simulated attacks. Our results show that provincial networks are more robust than the national network, suggesting that cross-regional collaboration fosters industrial clustering but also increases vulnerability to external shocks. Finally, we design a composite node importance ranking method that integrates multiple indicators for high accuracy. Furthermore, based on our findings, we offer policy implications to improve risk management. For example, to enhance network robustness, governments may consider promoting regional differentiation by encouraging diverse business roles across provinces. While our analysis focuses on China's UAV industry, the framework linking business similarity, multi-layer networks, and robustness can be applied to other interdependent high-tech industries, providing a basis for comparative studies.

The remainder of this article proceeds as follows. Section 2 elaborates data and models used in this work. Section 3 presents the main results. Section 4 discusses the findings, and Section 5 concludes.

2 Data and methods

2.1 Data sources

The primary dataset utilized in this study was obtained from Qichacha (https://www.qichacha.com/), a commercial data platform providing detailed firm-level information on Chinese enterprises. Our sample includes 2,120 listed UAV enterprises, with data covering corporate names, registered provinces, industrial classifications, business scopes, and other relevant operational characteristics. Non-listed and smaller-scale enterprises were systematically excluded due to the often limited availability and lower accuracy of operational data. Moreover, these enterprises generally contribute less to the core UAV industrial chain dynamics. Six provinces, namely, Guangdong, Jiangsu, Beijing, Zhejiang, Shandong, and Henan, were selected as representative regions, collectively constituting 60% of the total sample, as shown in Table 1. The remaining 40% of UAV enterprises are distributed across 28 other provincial-level administrative regions, which were

10.3389/fphy.2025.1651460

TABLE 1 Distribution of UAV enterprises by industry chain tier and region.

Region	Upstream	Midstream	Downstream	Overall	Proportion
China	398	1,548	1,452	2,120	—
Guangdong	126	254	290	434	20.47%
Jiangsu	57	179	168	236	11.13%
Beijing	50	136	136	193	9.10%
Zhejiang	45	119	130	169	7.97%
Shandong	29	85	86	123	5.80%
Henan	19	75	66	106	5.00%

not included in the regional comparison due to their smaller sample sizes. In this study, we use business similarity among UAV enterprises as a proxy for inter-firm connectivity instead of traditional trade data, as the latter reflects only short-term realized transactional activities rather than potential structural connections.

All observed UAV enterprises were manually classified into distinct segments of the industrial chain based on their business scopes. The UAV industry in China encompasses research and development, manufacturing, assembly, and distribution [33, 34]. The upstream segment primarily consists of raw materials, including metals and composites. The midstream involves the manufacturing of components such as engines, sensors, and avionics, while the downstream focuses on the assembly of UAVs and their distribution to end-users [35]. Notably, certain enterprises exhibit cross-tier operations spanning the entire industrial chain. Table 1 presents the national and provincial-level data for the UAV enterprise network.

2.2 Methodology

2.2.1 Construction of multi-layer UAV enterprise network model

The UAV enterprise network is modeled as a three-layer adjacency matrix $S = (A, O)$, where $A = \{A^{[1]}, A^{[2]}, A^{[3]}\}$ denotes intra-layer adjacency matrices representing connections within each industrial segment, while O captures inter-layer relationships. Specifically, $A^{[1]}$ represents the upstream network of the UAV industry chain, $A^{[2]}$ denotes the midstream network, and $A^{[3]}$ corresponds to the downstream network. Enterprises serve as network nodes, and each enterprise may appear in one or more layers, depending on whether its business scope includes activities related to that stage. Edges between enterprises within the same layer are established based on business similarity, quantified using the cosine similarity between their full registered business scope texts. This text-based measure emphasizes functional similarities that drive structural patterns within the network, independent of firm-level attributes such as ownership, location, or size. Prior to computation, all texts are first standardized by removing punctuation and stopwords, and vectorized using the term frequency-inverse document frequency (TF-IDF) method. The cosine similarity between enterprise i and j is calculated in Equation 1:

$$Similarity_{i,j} = \frac{\vec{T}_i \cdot \vec{T}_j}{\left|\vec{T}_i\right| \cdot \left|\vec{T}_j\right|} \tag{1}$$

where $\vec{T}_i$ and $\vec{T}_j$ denote the TF-IDF vectors of enterprise i and j's full business scope text. This enables quantification of operational proximity based on the actual semantic content of business scope. For example, a high similarity score between a firm focused on UAV R&D and flight control systems and another emphasizing UAV design and intelligent flight systems indicates functional overlap. All non-zero similarity values were retained as network edges to preserve the full spectrum of potential functional relationships.

Formally, the intra-layer connectivity within the θ-th layer is formally structured as $A^{[\theta]} = \left(Q^{[\theta]}, E^{[\theta]}, W^{[\theta]}\right)$, where $Q^{[\theta]}$ represents the set of enterprise nodes in the θ-th UAV industry layer, $E^{[\theta]}$ denotes the edge set of UAV in the θ-th layer, and $W^{[\theta]}$ specifies the weight of the θ-th layer network edge. If the similarity between enterprise i and enterprise j is non-zero, an edge $e_{i,j}^{[\theta]}$ is established with edge weight $W_{i,j}^{[\theta]}$. Otherwise, no edge connection is created between the two enterprises. Similarly, inter-layer relationships are defined as $O^{[\theta,\mu]} = \left(Q^{[\theta]}, Q^{[\mu]}, E^{[\theta,\mu]}, W^{[\theta,\mu]}\right)$. If a similarity relationship exists between enterprise i in layer θ and enterprise j in layer μ, an edge $e_{i,j}^{[\theta,\mu]}$ is added with a weight $W_{i,j}^{[\theta,\mu]}$. Notably, while the network is multilayered, this study focuses on intra-layer dynamics, as business similarity is mostly layer-specific. Vertical edges reflect enterprise presence across layers rather than strong functional interdependence. The three-layer UAV industry chain network ($n = 3$) formalized in Equation 2, with more network relationships detailed in Equation 3. Figure 1 provides a visual representation of this three-tier structure, including both intra-layer andinter-layer edges.

$$S = \begin{bmatrix} A^{[1]} & O^{[1,2]} & 0 \\ O^{[2,1]} & A^{[2]} & O^{[2,3]} \\ 0 & O^{[3,2]} & A^{[3]} \end{bmatrix} \tag{2}$$

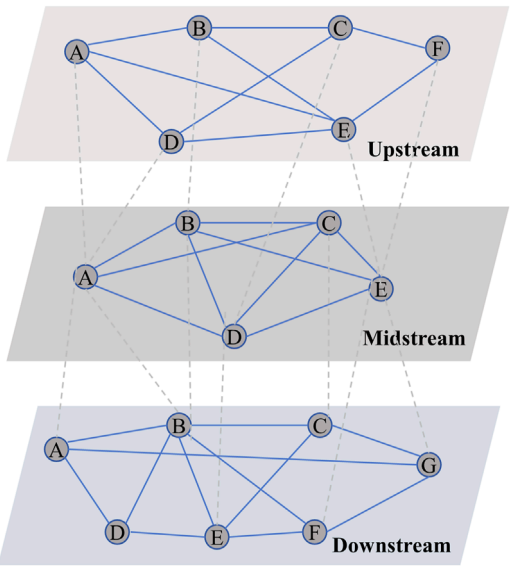


FIGURE 1
Three-tier UAV industry chain network based on business similarity. The circles represent nodes reflecting UAV enterprises, while the blue solid lines represent intra-layer edges derived from the similarity of business scopes. The gray dashed lines represent inter-layer edges, which exist when enterprises operate across layers or exhibit business similarity with enterprises in other layers.

$$S = \begin{bmatrix} \begin{pmatrix} W^{1}_{1,1} & \cdots & W^{1}_{1,n} \\ \vdots & \ddots & \vdots \\ W^{1}_{n,1} & \cdots & W^{1}_{n,n} \end{pmatrix} & \begin{pmatrix} W^{1,2}_{1,1} & \cdots & W^{1,2}_{1,n} \\ \vdots & \ddots & \vdots \\ W^{1,2}_{n,1} & \cdots & W^{1,2}_{n,n} \end{pmatrix} & 0 \\ \begin{pmatrix} W^{2,1}_{1,1} & \cdots & W^{2,1}_{1,n} \\ \vdots & \ddots & \vdots \\ W^{2,1}_{n,1} & \cdots & W^{2,1}_{n,n} \end{pmatrix} & \begin{pmatrix} W^{2}_{1,1} & \cdots & W^{2}_{1,1} \\ \vdots & \ddots & \vdots \\ W^{2}_{1,1} & \cdots & W^{2}_{1,1} \end{pmatrix} & \begin{pmatrix} W^{2,3}_{1,1} & \cdots & W^{2,3}_{1,n} \\ \vdots & \ddots & \vdots \\ W^{2,3}_{n,1} & \cdots & W^{2,3}_{n,n} \end{pmatrix} \\ 0 & \begin{pmatrix} W^{3,2}_{1,1} & \cdots & W^{3,2}_{1,n} \\ \vdots & \ddots & \vdots \\ W^{3,2}_{n,1} & \cdots & W^{3,2}_{n,n} \end{pmatrix} & \begin{pmatrix} W^{3}_{1,1} & \cdots & W^{3}_{1,1} \\ \vdots & \ddots & \vdots \\ W^{3}_{1,1} & \cdots & W^{3}_{1,1} \end{pmatrix} \end{bmatrix} \tag{3}$$

2.2.2 Measurement of structural robustness under normal conditions

The UAV enterprise network is defined by its nodes and edges, whose structural features are key to revealing business homogeneity among enterprises. The network's internal robustness is assessed through topological properties such as node connectivity and connection distribution, calculated using Gephi 0.10.1.

Degree is a fundamental parameter characterizing the local connectivity of nodes in the networks [36]. It quantifies the number of edges directly connected to a node, denoted as k_i. As formalized in Equation 4, for nodes i and j within a network of total size N, the adjacency matrix element a_{ij} indicates the presence (a_{ij} = 1) or absence (a_{ij} = 0) of a direct connection. Nodes with higher degrees are more central, thus reflecting their enhanced influence on structural robustness [36]. It is calculated by

$$k_i = \sum_{j=1}^{N} a_{ij} \tag{4}$$

Degree distribution represents the probability distribution of node degrees across the entire network. Defined as $P(k)$, the degree distribution reflects the fraction of nodes N_k with degree k relative to the network size N, as shown in Equation 5. The distribution of nodes can be determined based on the variation in the $P(k)$ function, and provides insights into the connectivity and structural prominence within the UAV enterprise networks.

$$P(k) = N_k / N \tag{5}$$

The clustering coefficient characterizes the transitivity of a network, representing the average likelihood that two neighboring nodes of a given local node are also interconnected [37]. Originally proposed by Watts and Strogatz, this metric gauges the tendency of nodes to form clusters within a network [37]. A high clustering coefficient indicates a strong clustering structure, whereas a low clustering coefficient suggests the absence or weakness of such a structure. Denoted as C, it is computed in Equation 6:

$$C = \frac{1}{N} \sum_{i=1}^{N} \frac{2N_i}{k_i(k_i - 1)} \tag{6}$$

Where N_i denotes the number of edges associated with node i, and k_i is the degree of node i. The clustering coefficient satisfies $0 \le C \le 1$. When $C = 0$, nodes remain unconnected, whereas $C = 1$ indicates a fully interconnected network.

The average path length is a global property representing the average value of all shortest paths between any two pairs of nodes [37]. It quantifies the overall efficiency and connectivity in complex networks, particularly in small-world networks and scale-free networks [37]. For a network with N nodes, the shortest path length between node i and j is denoted as d_{ij}. The average path length L is defined in Equation 7.

$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij} \tag{7}$$

2.2.3 Evaluation of structural robustness under targeted attacks

2.2.3.1 Methods of node importance ranking

To assess the structural robustness of the network under targeted attacks, it is essential to determine the node removal sequence by ranking node importance. Prior studies have typically employed single-index centrality measures such as degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality, each capturing a different aspect of node importance.

Degree centrality represents the most straightforward method for quantifying node centrality in network analysis [38]. Nodes with a high degree of centrality are deemed influential due to their extensive network connections, enabling efficient information propagation and substantial systemic influence. However, it overlooks a node's global impact in the network. It is calculated as Equation 8.

$$DC_i = \frac{k_i}{N-1} \tag{8}$$

Betweenness centrality identifies nodes that frequently lie on the shortest paths between other nodes, reflecting control over information flow [39]. As shown in Equation 9. The betweenness centrality BC_i of node i is defined as the ratio of the number of shortest paths between any pair of nodes j and l that pass through

node i ($N_{jl}(i)$) to the total number of all shortest paths between nodes j and nodes l (N_{jl}).

$$BC_i = \sum_{j \neq i \neq l} \frac{N_{jl}(i)}{N_{jl}} \tag{9}$$

Closeness centrality quantifies a node' centrality by measuring the average shortest path distance from the node to all other nodes [40]. Nodes with higher values are more central. The formula is presented as Equation 10, where CC_i denotes the closeness centrality of node i, and d_{ij} represents the shortest path distance between nodes i and j.

$$CC_i = \frac{1}{\sum_{j=1}^{N} d_{ij}} \tag{10}$$

Eigenvector centrality considers not only the quantity but also the importance of a node's neighbors, emphasizing influence within high-status neighborhoods [41]. This recursive relationship ensures that a node gains higher centrality if it is connected to other highly central nodes, as formalized in Equation 11, where c is a proportionality constant, and x_j denotes the eigenvector centrality of node j.

$$EC_i = c\sum_{j=1}^{N} a_{ij} x_j \tag{11}$$

While widely adopted, single centrality metrics may fail to capture the multidimensional nature of node influence in complex networks. To address this limitation, we propose a composite node importance index that incorporate richer structural information based on information-weighted assignment. Traditional weighting methods often rely on correlation-based ranking or similarity to ideal solution, which may overlook data complexity. In contrast, the information-weighted method is grounded in mathematical derivation, and quantifies variability and informational content across criteria. Specifically, we integrates three objective approaches: the CRITIC for contrast analysis, the Entropy Method for uncertainty, and Entropy-Weighted TOPSIS for ideal solution evaluation. To address scale and dimensionality disparities among indicator values, we implement min-max normalization to standardize these metrics within a unified range of 0–1. The normalized value x^* is calculated in Equation 12:

$$x* = (x - x_{\min})/(x_{\max} - x_{\min}) \tag{12}$$

where x is the actual value, and $x_{\max}$, $x_{\min}$ are the maximum and minimum value, respectively. Each method produces a weight vector for the four centrality indicators. These are assembled into a method-indicator matrix w_{ij}, where i = 1,2,3 denotes the method and j = 1,2,3,4 denotes the indicator. To quantify the informational contribution of each method, we apply an entropy-based fusion strategy. For each method i, the normalized weight distribution E_i across indicators in Equation 13:

$$E_i = -k\sum_{j=1}^{4} p_{ij} ln(p_{ij}), \text{where} p_{ij} = \frac{x*}{\sum_{i=1}^{3} x*}, \text{k} = 1/\ln(4) \tag{13}$$

The diversification degree, d_i = 1-E_i, is used to compute method-level weights in Equation 14, where E_i represents the entropy value associated with each method. Final indicator weight allocations are then determined as specified in Equation 15. Subsequently, the overall node importance score is then given by Equation 16.

$$\theta_i = \frac{d_i}{\sum_{i=1}^{3} d_i} \tag{14}$$

$$w_j = \sum_{i=1}^{3} \theta_i \cdot w_{ij} \tag{15}$$

$$NR_i = w_1 DC_i * + w_2 BC_i * + w_3 CC_i * + w_4 EC_i * \tag{16}$$

where NR_i denotes the comprehensive index of node importance, w_j represents the information weight derived from the three methods. DC_i^*, BC_i^*, CC_i^*, and EC_i^* represent the values of degree, betweenness, closeness, and eigenvector centrality for node i, respectively. The workflow of the proposed multi-indicators and multi-weight node ranking methodology is presented in Figure 2.

2.2.3.2 Network disintegration experiment

We conducted disintegration simulations on the three-layer UAV enterprise networks at both the national and six provincial levels to evaluate structural robustness. Specifically, nodes were removed according to the composite ranking method introduced earlier, in descending order of calculated importance. The simulations were performed with node removal rates from 0 to 1 in 0.1 increments. As the process was non-stochastic, repeated iterations were unnecessary. At each step, we calculated accessibility, connectivity, and efficiency to assess structural degradation. All simulations and visualizations, covering eight regions and six structural indicators, were implemented using Python 3.0.

Accessibility serves as an efficiency-oriented metric for industrial chain networks, quantifying the likelihood or ease with which any two nodes can communicate via alternative paths [42, 43]. It is measured by the average shortest path length within the maximum connected subgraph of the network, and the shorter the path, the higher the accessibility. The average value of the shortest path of the industrial chain network determines the transmission speed of business information between enterprises. It can be quantified as the ratio of the average value of the shortest path of the maximum connected subgraph after the network is attacked to the average shortest path length of the original system.

Connectivity evaluates the stability of industrial chain networks, particularly the connectivity density among enterprises [43]. It is quantified by the remaining capacity of the network after node failures. The proportion of the number of nodes retained in the surviving connected subgraph relative to the number of nodes in the original network size constitutes a direct indicator of the network's vulnerability.

Network efficiency serves as a metric to evaluate resource allocation and operational effectiveness in industry chain networks, particularly reflecting the collaborative performance among nodes [42]. It can be conceptualized as the ratio of the global efficiency of the network post-attack to the global efficiency of the original system.

3 Results

3.1 The ranks of node importance in UAV network

To ensure the objectivity of node importance ranking, we first employed the CRITIC method, Entropy Method, and Entropy-weighting TOPSIS Method to assign weights to the

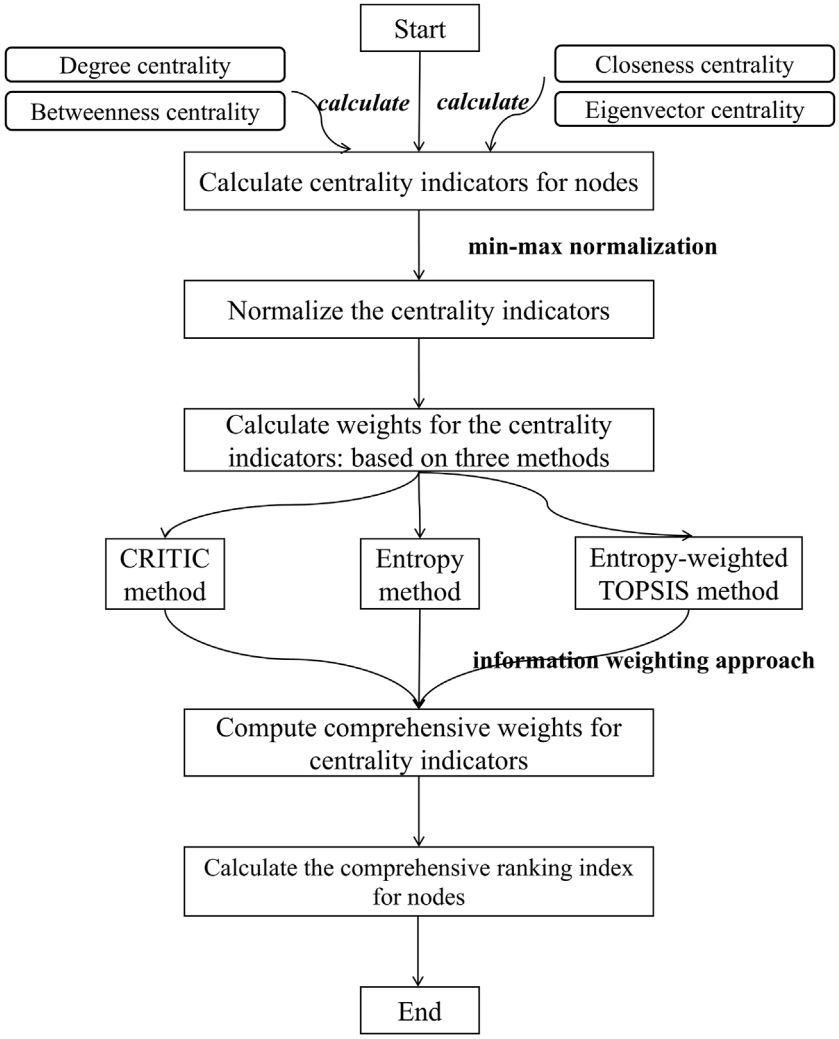


FIGURE 2
Workflow of the multi-indicators and multi-weight ranking method for node importance. The pipeline integrates four centrality metrics (degree, betweenness, closeness, and eigenvector centrality) through normalization, multi-method weight allocation (CRITIC, entropy, and entropy-weighted TOPSIS), and comprehensive weighted aggregation. The final ranking index synthesizes topological and information-theoretic perspectives to quantify node significance in complex networks.

four centrality metrics. Subsequently, we applied the information weighting method, a quantitative analysis approach suitable for complex systems and large-scale datasets, to reassign the weights. All the final information is presented in Table 2. Finally, based on the calculated weights, we ranked the importance of nodes at both the national and provincial levels. The detailed rankings can be found in Supplementary Table S1, with the top 30 nodes displayed due to space limitations. This distribution method enables a comprehensive evaluation of each node's importance within the network. By incorporating multiple centrality metrics, our approach achieves more objective, precise, and robust rankings for node importance in the UAV industry chain network.

To validate the methodology proposed in Section 2.2.3.2, we calculated node importance rankings in the UAV enterprise network and selected the top 30 nodes for further verification. Node information is provided in Supplementary Table S2, while network metrics are presented in Table 3. Compared to single-metric approaches, our method demonstrated significantly lower APL and variance, indicating higher connectivity and minimal structural changes under risk disturbances. It also exhibits the highest cohesion (0.94), resistance (0.92), and resilience (0.97), with relative improvements of 9.17%, 9.26%, and 4.29%, respectively. Although the absolute differences among the five metrics are not large, the overall performance consistently favors the information weight method. This trend is further supported by the resilience degradation curves (Figure 3), which show slower connectivity loss during progressive node removals. These findings confirm its enhanced robustness, superior ranking stability, and imporved network resilience under risk scenarios.

TABLE 2 Weight allocation of centrality indicators by different objective methods.

Method	Weight			
	Degree centrality	Betweenness centrality	Closeness centrality	Eigenvector centrality
CRITIC	21.15%	32.65%	21.99%	24.21%
Entropy Method	19.79%	36.82%	27.30%	16.09%
Entropy-weighting TOPSIS Method	31.38%	7.92%	28.55%	32.15%
Information weight	18.19%	47.96%	10.77%	23.09%

TABLE 3 Robustness test of node ranking methods based on network metrics.

Metrics	CRITIC	Entropy method	Entropy-weighting TOPSIS method	Information weight
Average path length	1.2046	1.3126	1.2667	1.1563
Cohesion	0.8935	0.8587	0.9239	0.9374
Resistance	0.8985	0.8437	0.8674	0.9218
Resilience	0.9289	0.9291	0.9615	0.9687
Variance	6.3073	5.9270	6.9298	5.4817

3.2 Structural robustness at the national-level

3.2.1 Systemic topological characteristics

Figure 4 presents the topological characteristics of the national-level UAV enterprise operation network, including node-level (i.e., degree distribution) and edge-level (i.e., clustering coefficient) indicators. The results show that the overall structural robustness at the national level is consistently below the provincial average, as evidenced by its higher average degree and clustering coefficient, along with a shorter average path length. This suggests that interprovincial collaboration enhances connectivity but may also increase systemic risk due to greater business homogeneity.

Figure 4A illustrates the cumulative degree distributions of UAV enterprise operations at the national level. The results reveal that the national UAV network follows a power-law distribution, with linear trends in log-log coordinates [44]. This implies a "rich-get-richer" structure, dominated by a few highly connected core enterprises, driven by preferential attachment [45, 46]. A similar pattern is observed across different industrial chain segments, where internal resources become increasingly concentrated among leading enterprises [47, 48]. This concentration may weaken the innovation capacity of small and medium-sized enterprises (SMEs) and raise systemic risk. Such a pattern likely stems from the high technological barriers inherent in the UAV industry, consistent with evidence that Matthew effects are more prominent in high-tech sectors [49, 50]. This self-reinforcing process suggests that enterprises with existing advantages are more likely to attract new collaborations, further amplifying their centrality and influence in the network.

Additional topological results, shown in Figure 4B, reveal uneven structural robustness across the UAV industry chain at the national level. The upstream UAV enterprise network exhibits the highest robustness, while the downstream segment is the most fragile. The downstream segment has a high average degree, strong clustering, and short average path length, indicating broad business reach and dense clusters of similar enterprises. This may be due to the downstream UAV sector's greater maturity, with widespread applications in agriculture, tourism, emergency services, and logistics. However, such diversification may intensify long-term systemic risk due to business homogenization. Despite its apparent structural density, the downstream network is more prone to rapid collapse under targeted attacks.

3.2.2 Simulation of targeted attack scenarios

The simulation results of the structural robustness of the national-level UAV enterprise network under targeted attacks are presented in Figure 5, with detailed data available in Supplementary Table S3. As nodes are progressively removed, the UAV enterprise network maintains fundamental structural robustness during the initial phase but undergoes rapid deterioration beyond a critical threshold (node removal rate >60%). This phenomenon is attributed to business homogeneity. Similarity-driven clustering enhances communication density within clusters but also increases systemic risk by fostering dependency on a limited number of key nodes with overlapping business characteristics. These findings align with previous studies; business similarity exacerbates systemic vulnerability by concentrating risks within core clusters [7, 10, 11].

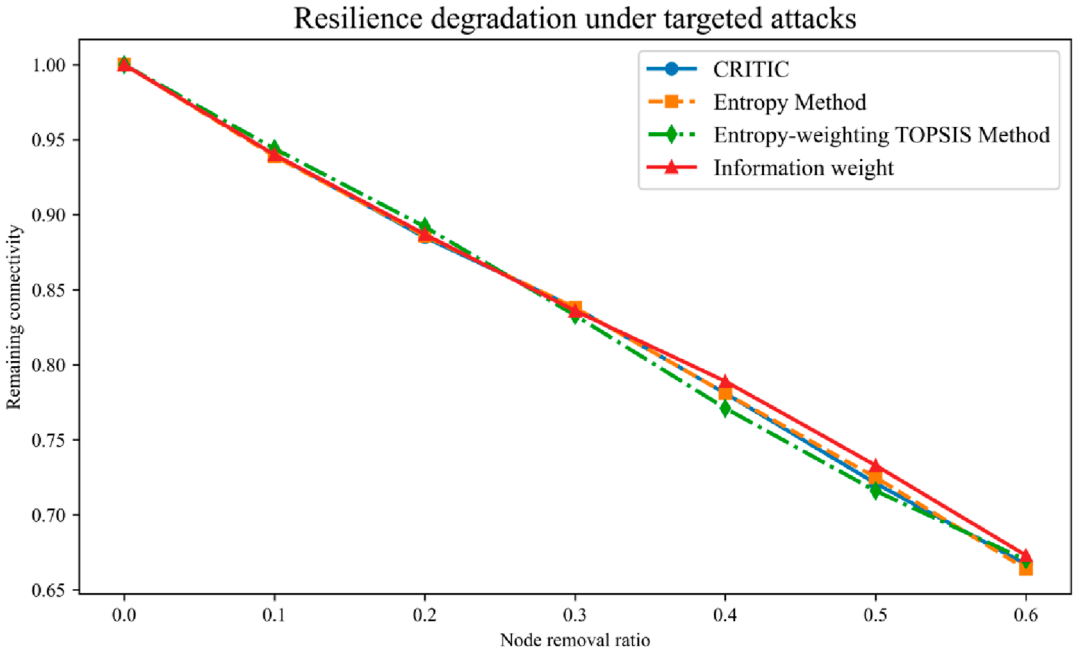


FIGURE 3
Resilience degradation of the UAV enterprise network under targeted attacks. The y-axis represents remaining connectivity (i.e., the largest connected component relative to the original network), and the x-axis indicates the node removal rate. Blue, orange, and green lines represent CRITIC, Entropy, and TOPSIS methods, while the red line is the proposed information weight method, which shows superior robustness beyond a 0.2 removal rate.

Specifically, as depicted in Figures 5a,b, APL remains relatively constant throughout most of the node removal process, while the accessibility of the national UAV network declines after approximately 60% node removal. This implies that while path lengths among connected nodes remain short, the overall connectedness decreases, isolating clusters of similar businesses. Furthermore, the sharp shrinkage of the maximal connected subgraph (Figure 5c) and the sharp decline in connectivity (Figure 5d) indicate the fragmentation and vulnerability arising from interdependencies of enterprises with overlapping business activities. Such fragmentation is considered to threaten the information and resource flows critical for maintaining network functionality. Nevertheless, global efficiency stays stable or slightly increases under high node removal (Figure 5e), while network efficiency in Figure 5f even exceeds 1 (up to 1.03). This unexpected trend is likely driven by tighter local clustering among surviving firms. To verify this, we compared network efficiency and clustering coefficients before and after node removal (Figure 6). Both metrics exhibit a nearly parallel upward trend, suggesting that the efficiency gain stems from denser intra-cluster ties rather than enhanced overall robustness.

3.3 Structural robustness at the provincial-level

3.3.1 Systemic topological characteristics

The cumulative degree distributions of provincial-level UAV enterprise networks are presented in Figure 7. Similar to the national network, all six provincial networks exhibit a power-law distribution, a hallmark of scale-free networks. This implies the presence of a few hub enterprises with extensive business overlaps, reinforcing the existence of hierarchical structures with uneven inter-firm connectivity. Such a pattern indicates that once an enterprise becomes central in a provincial network, its likelihood of forming new connections increases, consolidating its dominant position and reinforcing structural inequality across firms.

Further topological analysis in Figure 8 reveals pronounced regional differences. As shown in Figure 8a, Jiangsu demonstrates a high clustering coefficient, elevated average degree, and the lowest APL, suggesting a tightly connected structure with high business similarity among UAV enterprises. While this configuration enhances internal collaboration and operational efficiency, it also indicates increased systemic risk due to close interdependence. In contrast, Henan exhibits low clustering and connectivity, implying relatively independent business operations among enterprises, likely reflecting the province's less developed UAV industry base. Interestingly, although Guangdong holds the highest average degree, its network appears loosely clustered and topologically fragmented, indicating a broader but weaker pattern of inter-firm linkages.

Subnetwork analysis across different segments of the UAV industrial chain is shown in Figures 8b–g. A consistent pattern emerges; economically developed provinces such as Guangdong, Jiangsu, Beijing, and Zhejiang exhibit stronger industrial collaboration. Their networks show higher average degree and clustering coefficients across upstream, midstream, and downstream segments compared to less developed regions like Henan and Shandong, suggesting broader industrial coverage and tighter inter-enterprise relationships. While such

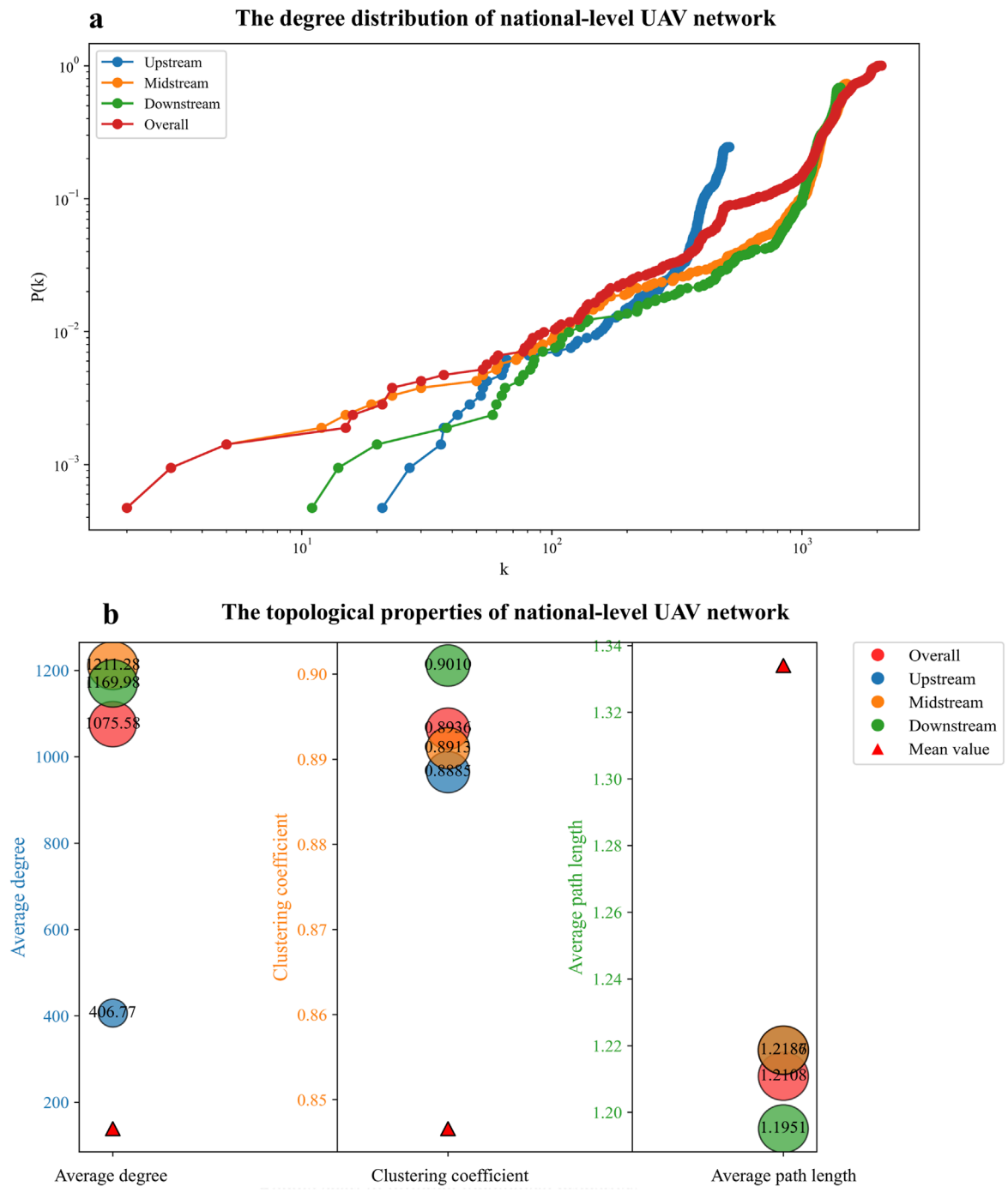


FIGURE 4
Topological characteristics of the national-level UAV enterprise network. Results are shown for upstream (blue), midstream (orange), downstream (green), and overall (red) segments. **(a)** The line chart presents the cumulative degree distribution. The horizontal axis (K) represents node degree, spanning 10^1 to 10^5, while the vertical axis ($P(K \geq (K))$) shows the probability of nodes with degrees $\geq k$, scaled from 10^{-3} to 10^3. **(b)** The bubble chart displays three structural indicators: degree, clustering coefficient, and average path length. Bubble colors represent industrial chain stages, and bubble sizes reflect the magnitude of the indicators. Red triangles denote the provincial average for each metric.

characteristics promote production efficiency and innovation through specialization and cooperation, they also contribute to concentrated vulnerability. Additionally, structural imbalances across industrial chain stages are evident in certain provinces. For instance, Henan's downstream segment has a significantly lower average degree (53.396) and clustering coefficient (0.7705) compared to its upstream and midstream segments, revealing weak business cohesion in application services. This suggests potential disruptions or fragmentation in the regional UAV value chain.

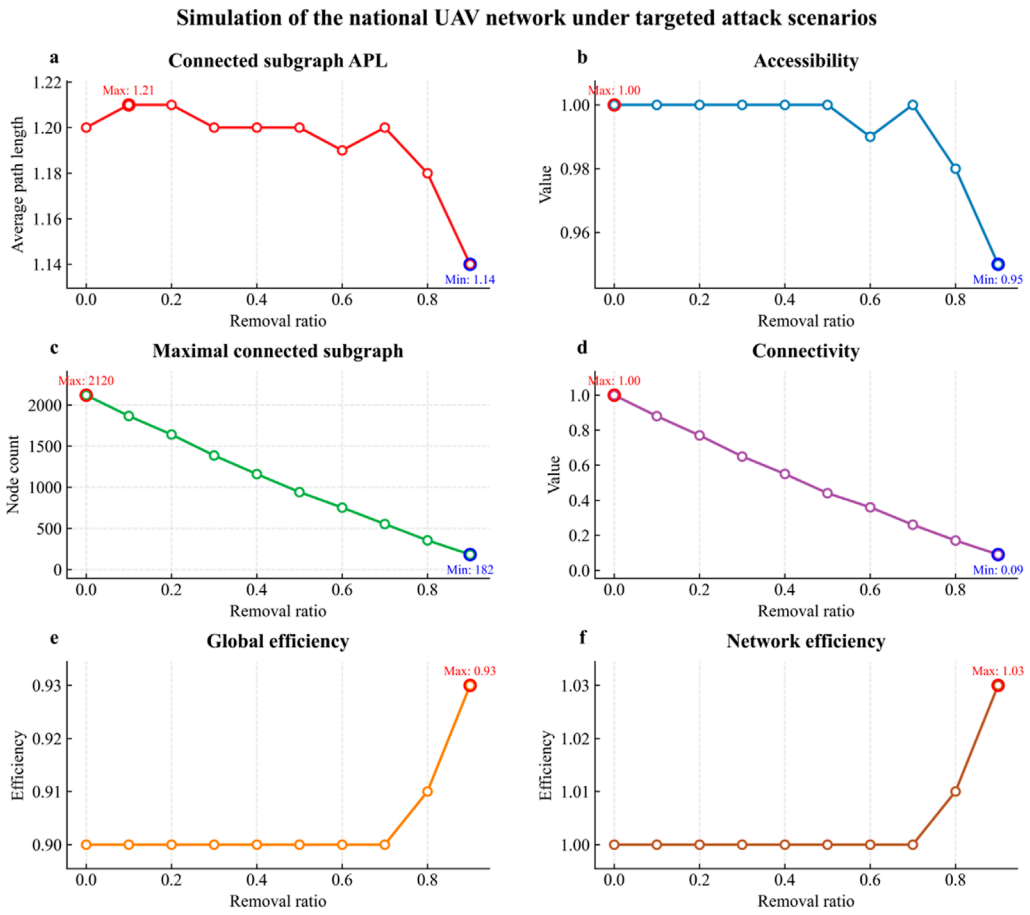


FIGURE 5
Simulation of the national UAV enterprise network under targeted attack scenarios. Nodes were sequentially removed based on the ranks of node importance in increments of 0.1 (from 0 to 0.9); results at the full removal ratio (1.0), where all indicators drop to zero, are omitted for clarity. Red circles mark peak values, and blue circles indicate minima. Evaluated metrics include: **(a)** average path length (APL) of the largest connected subgraph, **(b)** accessibility, **(c)** size of the maximal connected subgraph, **(d)** connectivity, **(e)** global efficiency, and **(f)** network efficiency.

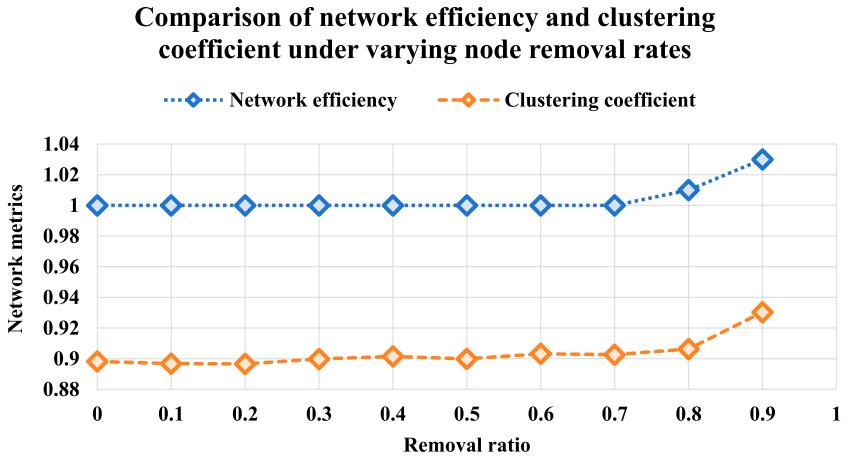


FIGURE 6
Comparison of network efficiency and clustering coefficient under varying node removal rates. In the figure, the network efficiency is depicted by the blue line, whereas the clustering coefficient is shown by the orange line.

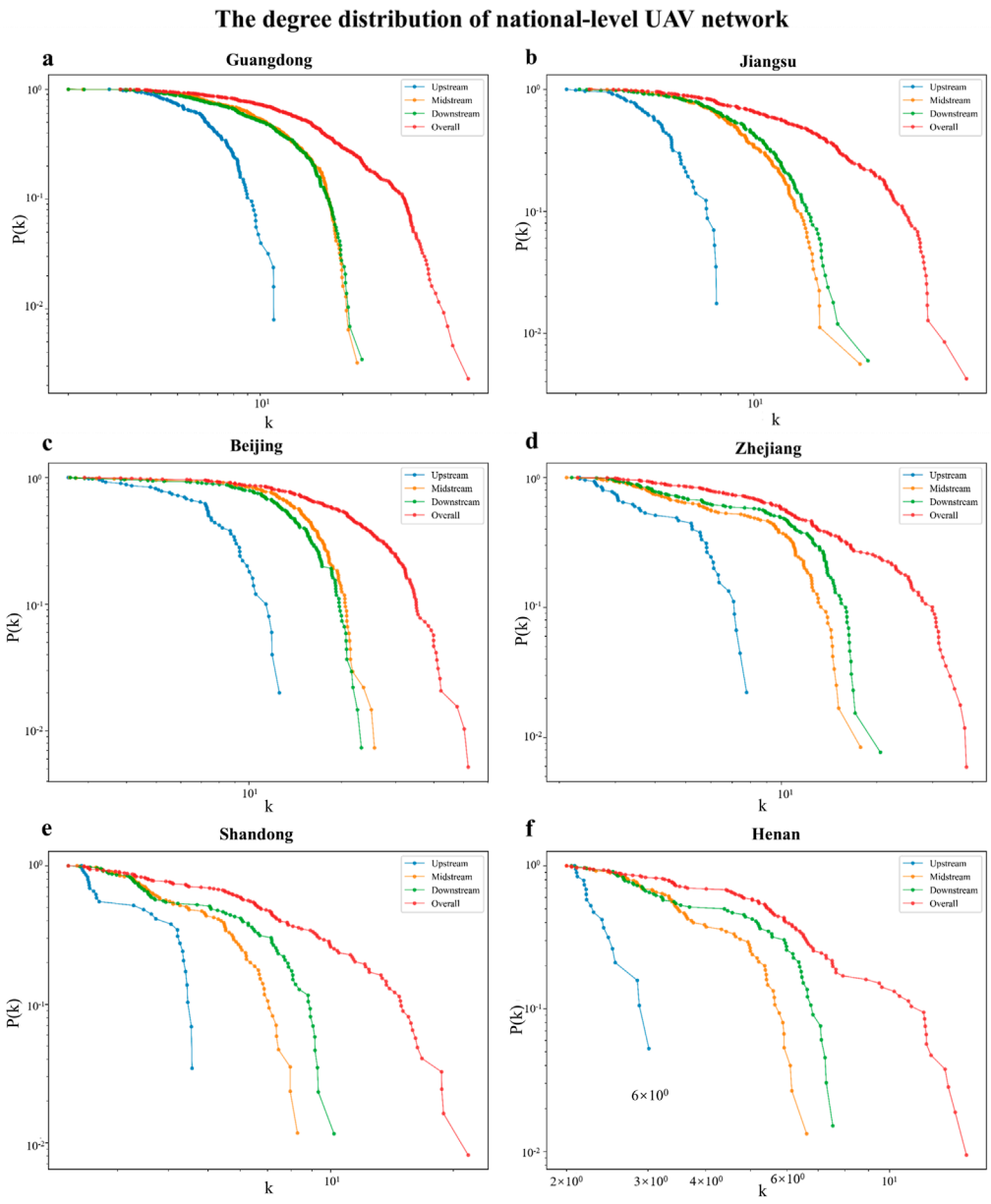


FIGURE 7
The cumulative degree distribution of provincial-level UAV enterprise network. Six sub-figures compare the cumulative degree distributions of **(a)** Guangdong, **(b)** Jiangsu, **(c)** Beijing, **(d)** Zhejiang, **(e)** Shandong, and **(f)** Henan provinces across upstream (blue), midstream (green), downstream (orange), and overall (red) segments. The horizontal axis (k) represents node degree, while the vertical axis ($P(K \geq k)$) shows the probability of nodes with degrees greater than or equal to k.

3.3.2 Simulation of targeted attack scenarios

The simulation results of provincial UAV enterprise operation networks under targeted node removal are illustrated in Figure 9, with detailed numerical results available in Supplementary Table S4. The simulations reveal a consistent degradation pattern across all six provincial networks, as shown by a reduction in maximal connected subgraph size, network connectivity, and global efficiency. However, up to a 50% removal ratio, APL and accessibility remain relatively stable, indicating preserved local communication and operational buffering capacity despite growing global fragmentation. Based on a comparative evaluation of key robustness metrics, the networks can be categorized into four types, namely, efficiency-robustness coexisting networks (Jiangsu, Beijing), highly centralized but fragile networks (Guangdong), low-efficiency and low-robustness networks (Henan, Shandong), and moderately robust networks (Zhejiang).

In detail, Jiangsu and Beijing exhibit strong efficiency-robustness coexistence. In Jiangsu, APL and accessibility remain virtually unchanged throughout the removal process (see Figures 9a,b), and the maximal connected subgraph retains a substantial size

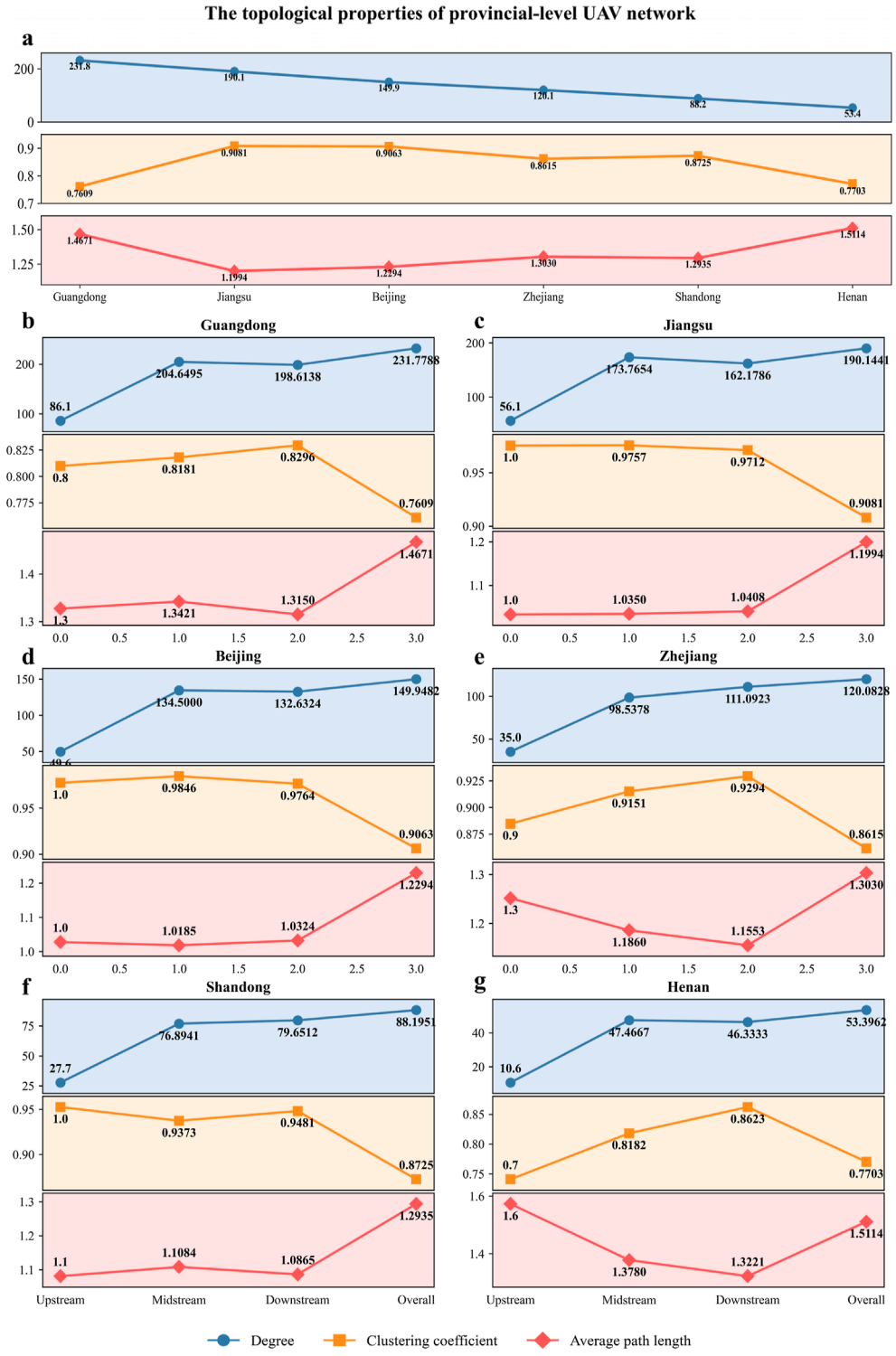


FIGURE 8
Topological characteristics of the provincial-level UAV enterprise network. The results are presented for the degree (circles), clustering coefficient (squares), and average path length (diamonds). Panels show **(a)** the overall network across six provinces, and the upstream, midstream, downstream, and overall segments of **(b)** Guangdong, **(c)** Jiangsu, **(d)** Beijing, **(e)** Zhejiang, **(f)** Shandong, and **(g)** Henan.

even under 70% node removal (see Figure 9c). Both global and network efficiency remain nearly constant (see Figures 9e,f), indicating an internally cohesive, well-connected structure with sufficient redundancy to withstand external perturbations. Similarly, Beijing shows minimal disruption in local structural indicators and maintains network efficiency close to 1 across all stages, despite a

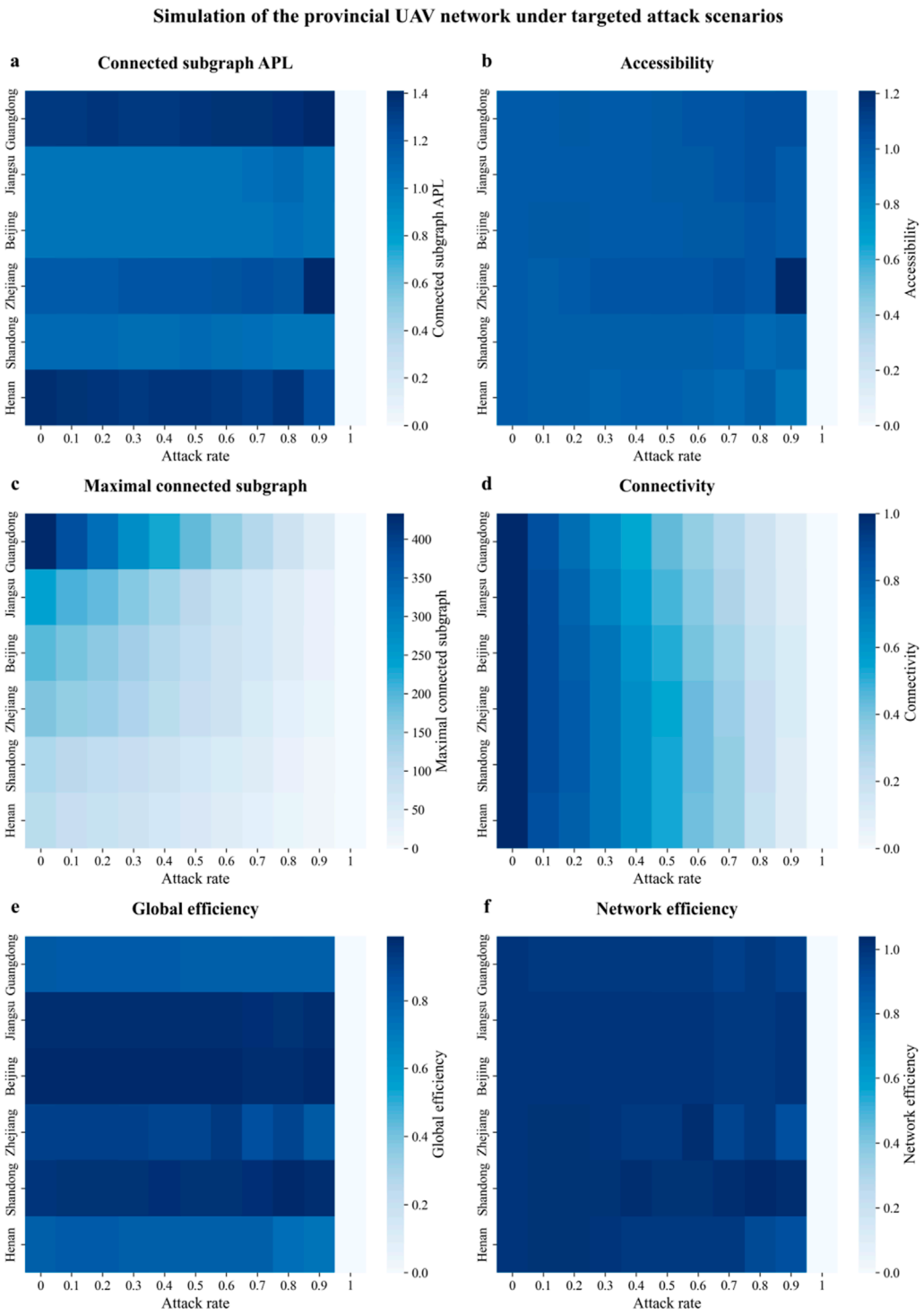


FIGURE 9
Simulation results of the provincial-level UAV enterprise network under targeted attack. The results cover six representative provinces, including Guangdong, Jiangsu, Beijing, Zhejiang, Shandong, and Henan. The heatmaps represent the values of six topological resilience metrics, where darker shades indicate higher values. Panels show **(a)** Connected subgraph APL, **(b)** Accessibility, **(c)** Max connected subgraph, **(d)** Connectivity, **(e)** Global efficiency, and **(f)** Network efficiency.

slightly faster degradation in connectivity (see Figures 9d,f). These observations are consistent with the Robustness-efficiency tradeoff framework (RETO) in complex network analysis [51, 52].

Guangdong represents a highly centralized but fragile structure. As shown in Figure 8c, it possesses the most extensive initial connectivity (433 nodes in the maximal subgraph), and its structural

integrity deteriorates rapidly once nodes are removed. APL and accessibility rise sharply (see Figures 9a,b), and connectivity drops to nearly zero at a 90% removal rate (see Figure 9d), indicating a lack of sufficient redundancy. Despite its reach, the network's centralized and loosely clustered structure makes it particularly vulnerable to targeted attacks.

Henan and Shandong display low efficiency and low robustness. Henan begins with the smallest subgraph, the highest APL, and the lowest global efficiency in Figures 9a,c,e, and suffers rapid and near-complete structural collapse at higher attack levels. Shandong shows similar fragility in connectivity (see Figure 9d), although global and network efficiency remain relatively stable or slightly improved (see Figures 9e,f), possibly due to peripheral node elimination and the persistence of a small core cluster. Nonetheless, both networks are characterized by weak inter-firm linkages and fragmented topologies.

Zhejiang occupies an intermediate position in terms of structural robustness. Although its APL and global efficiency begin to decline more noticeably in the later stages of attack (Figures 9a,e), and connectivity drops to 0.12 at the 90% removal level (Figure 9d), the overall network shows better structural stability compared to Henan and Shandong. This suggests that Zhejiang's network is denser but still partially reliant on central nodes, leading to a moderate level of vulnerability.

4 Discussion

4.1 Result analysis

Industrial clusters are a specific type of industrial organization, where enterprises with similar business scopes tend to co-locate and form tightly-knit networks [7, 27, 28]. Driven by business homogeneity, such clustering not only shapes the internal connectivity patterns of enterprise networks but also increases their vulnerability to external disruptions [7, 53]. While local collaboration within these clusters can improve operational efficiency, it simultaneously heightens systemic risk by concentrating dependencies on a few central enterprises [10, 11]. These risks are particularly pronounced under targeted disruptions, where the failure of key nodes can trigger widespread structural fragmentation [54, 55].

Our study applies a complex network model based on business similarity and network disintegration experiments to investigate the structural robustness of UAV enterprise networks at national and provincial levels. The following observations are obtained from the study. First, both levels exhibit power-law degree distributions, indicating a typical scale-free structure. This suggests that a few enterprises with high business similarity play disproportionately central roles in network connectivity, a pattern consistent with the Matthew effect in industrial development [49, 50]. Second, the national network is less robust than the provincial average, with inter-provincial collaboration increasing systemic risk despite enhancing connectivity. Notably, robustness varies across supply chain stages, with upstream segments being the most resilient and downstream segments the most vulnerable. This discrepancy is likely due to higher business similarity in downstream sectors, which increases mutual dependencies. Third, our simulation experiments indicate that the national UAV enterprise network can tolerate targeted attacks involving up to 60% of central nodes before structural collapse becomes irreversible. Fourth, regional heterogeneity is observed, with economically developed regions such as Guangdong, Jiangsu, Beijing, and Zhejiang demonstrate stronger internal coordination. Finally, based on network disintegration patterns, provincial networks can be categorized into four structural types, namely, resilient and efficient, centralized but fragile, weakly connected and inefficient, and moderately robust. These findings provide insights into how business similarity, clustering, and network topology jointly shape the robustness of emerging industrial systems such as UAVs.

4.2 Policy implications

Building on the insights from our analysis, we reexamine the structural vulnerabilities of China's UAV enterprise network under both normal conditions and targeted attacks. Accordingly, we propose actionable strategic recommendations at a comprehensive level for policymakers, illustrated as follows.

First and foremost, national policy should prioritize inter-provincial coordination, particularly between midstream and downstream enterprises. Our analysis conenterprises that the national UAV enterprise network, with extensive cross-regional linkages, has higher overall communication efficiency than the average provincial network. This aligns with previous studies suggesting that diversified inter-organizational ties boost adaptive capacity in technological networks [37, 56]. However, we also find that vulnerability concentrates in downstream segments, which are prone to cascading failures under targeted attacks. This fragility stems in part from the high business similarity among downstream enterprises, leading to the emergence of highly central nodes. To address this, the vertical linkages between midstream and downstream enterprises should be promoted to reduce dependency on key nodes. This can be achieved through coordinated industrial planning with cross-segment and cross-regional collaboration, especially by fostering partnerships among enterprises with complementary rather than homogeneous business scopes. Such interventions are in line with modular network design principles [57], which emphasize decentralization and functional compartmentalization. To support the implementation of these recommendations, policy tools such as the publication of industrial chain maps and subsidies for cross-regional collaboration should be considered.

Second, local governments should actively promote diversified and complementary industrial collaboration to enhance resilience against external shocks. Our provincial-level network analysis reveals that economically developed regions such as Guangdong, Jiangsu, Beijing, and Zhejiang exhibit weaker structural robustness. However, this is primarily driven by over-concentration around a few central enterprises, which increases systemic vulnerability. This pattern is consistent with empirical evidence showing that industrial clustering increases dependency on core nodes, thus increasing network fragility [7, 27]. Therefore, local governments in these regions could encourage heterogeneous business linkages that go beyond enterprises with similar scopes to decentralize critical hubs and improve structural robustness. In contrast, less developed provinces like Henan and Shandong tend to

have fragmented and inefficient industrial networks, limiting their capacity for coordinated growth. Here, fostering diversified forms of collaboration can accelerate the formation of a resilient industrial ecosystem. This recommendation is grounded in research on regional innovation systems [56, 58, 59], which emphasizes that balanced network topologies improve adaptive capacity and systemic stability.

Third, category-based differentiated regional policies and interventions are essential for addressing the structural heterogeneity of China's UAV enterprise networks. Our typological analysis classifies provincial networks into four distinct categories based on their efficiency and robustness. This typology provides a valuable foundation for targeted policy intervention, as emphasized in the literature on regional development asymmetry and adaptive industrial policy [60, 61]. In centralized but fragile regions like Guangdong, policy efforts should prioritize decentralization strategies, such as supporting secondary hubs and fostering horizontal collaborations among enterprises with complementary business scopes to reduce similarity-driven clustering. While structural upgrading is critical for inefficient and fragile regions such as Henan and Shandong. Governments in these areas should expedite infrastructure development, enhance innovation capacity, and encourage business diversification. Contrastingly, regions already displaying strong structural robustness should focus on sustaining current advantages while preparing for demand shocks and technological shifts. If well-designed and consistently implemented, these differentiated strategies can significantly enhance the structural robustness of China's UAV sector by reducing uneven regional development.

5 Conclusion

This study investigates the structural robustness of China's UAV enterprise operation networks using complex network models based on business similarity. Our analysis reveals that both national- and provincial-level networks exhibit scale-free characteristics, with robustness varying significantly across regions and supply chain segments. Notably, downstream enterprises, often clustered due to high business similarity, are particularly vulnerable to targeted disruptions. At the provincial level, economically developed regions tend to be more fragile due to over-concentration. We further identify four distinct types of provincial network structures, reflecting varying degrees of efficiency and robustness. These findings underscore the dual role of business similarity in fostering local coordination while also amplifying systemic risk through structural homogeneity. Policy implications highlight the importance of inter-provincial coordination, complementary inter-firm linkages, and region-specific strategies in enhancing the robustness of emerging industrial systems.

Admittedly, this work suffers from three notable limitations. First, despite its size, the Qichacha database could leave out new or unregistered UAV enterprises, which may bias the sample and underrepresent informal network users. Second, although excluding non-listed enterprises ensures data reliability, it may overlook peripheral participants that contribute to the system's redundancy and flexibility in actual networks. Third, the analysis does not further decompose the factors contributing to business similarity or examine their differentiated effects on network robustness. These restrictions might lead to an overestimation of network vulnerability in some regions and an underidentification of adaptive structures. Future research should leverage more comprehensive datasets and investigate the heterogeneous drivers of similarity across regions to support more targeted policy interventions.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.qichacha.com/.

Author contributions

XZ: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – original draft, Writing – review and editing. JG: Data curation, Project administration, Resources, Supervision, Writing – review and editing. SL: Conceptualization, Project administration, Resources, Supervision, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1651460/full#supplementary-material

References

1. GLOBE NEWSWIRE. Unmanned aerial Vehicle (UAV) Drones Market Size expected to Reach USD 169.7 Bn by 2033 (2024). Available online at: https://www.globenewswire.com/news-release/2024/08/09/2927741/0/en/Unmanned-Aerial-Vehicle-UAV-Drones-Market-Size-Expected-to-Reach-USD-169-7-Bn-by-2033.html (Accessed March 1, 2025).

2. Magistretti S, Dell'Era C. Unveiling opportunities afforded by emerging technologies: evidences from the drone industry. *Technol Anal Strateg Manag* (2019) 31(5):606–23. doi:10.1080/09537325.2018.1538497

3. Nex F, Armenakis C, Cramer M, Cucci DA, Gerke M, Honkavaara E, et al. UAV in the advent of the twenties: where we stand and what is next. *SPRS J Photogramm Remote Sens* (2022) 184:215–42. doi:10.1016/j.isprsjprs.2021.12.006

4. Marshall A. *Principles of economics*. London: Macmillan (1920).

5. Krugman P. The role of geography in development. *Int Reg Sci Rev* (1994) 16(1):48–62. doi:10.1177/016001769401600106

6. Porter ME. Clusters and the new economics of competition. *Harv Bus Rev* (1998) 76:77–90.

7. Becattini G. The Marshallian industrial district as a socio-economic notion. *Rev Écon Ind* (2017) 157(1):13–32. doi:10.4000/rei.6507

8. DiMaggio PJ, Powell WW. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. *Am Sociol Rev* (1983) 48(2):147–60. doi:10.2307/2095101

9. Ye D, Zheng L, He P. Industry cluster innovation upgrading and knowledge evolution: a simulation analysis based on small-world networks. *SAGE Open* (2021) 11(4):21582440211031604. doi:10.1177/21582440211031604

10. Humphrey J, Schmitz H. How does insertion in global value chains affect upgrading in industrial clusters? *Reg Stud* (2002) 36(10):1017–27. doi:10.1080/0034340022000022198

11. Gao J, Zhang W, Guan T, Feng Q, Mardani A. The effect of manufacturing agent heterogeneity on enterprise innovation performance and competitive advantage in the era of digital transformation. *J Bus Res* (2023) 155:113387. doi:10.1016/j.jbusres.2022.113387

12. Markets and Data. China unmanned aerial vehicle market assessment, opportunities and forecast, 2017-2031F (2024). Available online at: https://www.marketsandata.com/industry-reports/china-unmanned-aerial-vehicle-market (Accessed March 1, 2025).

13. Civil Aviation Administration of China. *China civil unmanned aerial vehicle development report 2023-2024* (2024).

14. Zhou F, Wen H. Trade policy uncertainty, development strategy, and export behavior: evidence from listed industrial companies in China. *J Asian Econ* (2022) 82:101528. doi:10.1016/j.asieco.2022.101528

15. Jahani H, Khosravi Y, Kargar B, Ong KL, Arisian S. Exploring the role of drones and UAVs in logistics and supply chain management: a novel text-based literature review. *Int J Prod Res* (2024) 63:1873–97. doi:10.1080/00207543.2024.2373425

16. Al-Dosari K, Deif AM, Kucukvar M, Onat N, Fetais N. Security supply chain using UAVS: validation and development of a UAV-based model for Qatar's mega sporting events. *Drones* (2023) 7(9):555. doi:10.3390/drones7090555

17. Cheng AL, Fuchs ERH, Karplus VJ, Michalek JJ. Electric vehicle battery chemistry affects supply chain disruption vulnerabilities. *Nat Commun* (2024) 15:2143. doi:10.1038/s41467-024-46418-1

18. Scholten K, Sharkey Scott P, Fynes B. Benefits and constraints of temporary sourcing diversification in supply chain disruption and recovery. *Int J Phys Distrib Logist Manag* (2020) 50(9/10):1019–41. doi:10.1108/IJPDLM-02-2020-0054

19. Bhattacharya B. A reliability based measure of structural robustness for coherent systems. *Struct Saf* (2021) 89:102050. doi:10.1016/j.strusafe.2020.102050

20. Dalton SK, Atamturktur S, Farajpour I, Juang CH. An optimization based approach for structural design considering safety, robustness, and cost. *Eng Struct* (2013) 57:356–63. doi:10.1016/j.engstruct.2013.09.040

21. Starossek U, Haberland M. Approaches to measures of structural robustness. *Struct Infrastruct Eng* (2011) 7(7-8):625–31. doi:10.1080/15732479.2010.501562

22. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: structure and dynamics. *Phys Rep* (2006) 424(4-5):175–308. doi:10.1016/j.physrep.2005.10.009

23. Xu M, Deng W, Zhu Y, Lü L. Assessing and improving the structural robustness of global liner shipping system: a motif-based network science approach. *Reliab Eng Syst Saf* (2023) 240:109576. doi:10.1016/j.ress.2023.109576

24. Artime O, Grassia M, De Domenico M, Gleeson JP, Makse HA, Mangioni G, et al. Robustness and resilience of complex networks. *Nat Rev Phys* (2024) 6(2):114–31. doi:10.1038/s42254-023-00676-y

25. Bryan SS, Alex C. *Tall building structures: analysis and design*. Wiley (1991).

26. Feng Q, Zhao X, Fan D, Cai B, Liu Y, Ren Y. Resilience design method based on meta-structure: a case study of offshore wind farm. *Reliab Eng Syst Saf* (2019) 186:232–44. doi:10.1016/j.ress.2019.02.024

27. Delgado M, Porter ME, Stern S. Defining clusters of related industries. *J Econ Geogr* (2016) 16(1):1–38. doi:10.1093/jeg/lbv017

28. Liu J, Zhou M, Wang S, Liu P. A comparative study of network robustness measures. *Front Comput Sci* (2017) 11:568–84. doi:10.1007/s11704-016-6108-z

29. Chen Q, Wang K, Zhang Y, Guan Q, Xing J, Long T, et al. Simulation analysis on the evolution driven model of global copper ore trade network. *Front Phys* (2025) 13:1559799. doi:10.3389/fphy.2025.1559799

30. Xu J, Li J, Charles V, Zhao X. Evolution of the rare earth trade network: a perspective of dependency and competition. *Geosci Front* (2024) 15(3):101653. doi:10.1016/j.gsf.2023.101653

31. Zhang Y, Wang J, Song Y. Trade networks and corporate ESG performance: evidence from Chinese resource-based enterprises. *J Environ Manag* (2024) 367:122079. doi:10.1016/j.jenvman.2024.122079

32. Zhang YT, Zhou WX. Structural evolution of international crop trade networks. *Front Phys* (2022) 10:926764. doi:10.3389/fphy.2022.926764

33. U.S.-China Economic AND Security Review Commission. China's industrial and military robotics development (2016). Available online at: https://www.uscc.gov/sites/default/files/Research/DGI_China's%20Industrial%20and%20Military%20Robotics%20Development.pdf (Accessed March 31, 2025).

34. Chinadaily. Sky's the limit in UAV autonomous push (2024). Available online at: https://www.chinadailyhk.com/hk/article/586479 (Accessed March 31, 2025).

35. Wu Q, Su Y, Tan W, Zhan R, Liu J, Jiang L. UAV path planning trends from 2000 to 2024: a bibliometric analysis and visualization. *Drones* (2025) 9(2):128. doi:10.3390/drones9020128

36. Barabási AL, Albert R. Emergence of scaling in random networks. *Science* (1999) 286(5439):509–12. doi:10.1126/science.286.5439.50

37. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* (1998) 393(6684):440–2. doi:10.1038/30918

38. Evans TS, Chen B. Linking the network centrality measures closeness and degree. *Commun Phys* (2022) 5(1):172. doi:10.1038/s42005-022-00949-5

39. Freeman LC. A set of measures of centrality based on betweenness. *Sociom* (1977) 40(1):35–41. doi:10.2307/3033543

40. Freeman LC. Centrality in social networks: conceptual clarification. *Soc Netw* (1978) 1:215–39. doi:10.1016/0378-8733(78)90021-7

41. Bonacich P. Some unique properties of eigenvector centrality. *Soc Netw* (2007) 29(4):555–64. doi:10.1016/j.socnet.2007.04.002

42. Gutiérrez J, Monzon AJMP, Piñero JM. Accessibility, network efficiency, and transport infrastructure planning. *Environ Plan A* (1998) 30(8):1337–50. doi:10.1068/a30

43. Labi S, Faiz A, Saeed TU, Alabi BNT, Woldemariam W. Connectivity, accessibility, and mobility relationships in the context of low-volume road networks. *Transp Res Rec* (2019) 2673(12):717–27. doi:10.1177/0361198119854091

44. Newman ME. Power laws, pareto distributions and Zipf's law. *Contemp Phys* (2005) 46(5):323–51. doi:10.1080/00107510500052444

45. Powell WW, White DR, Koput KW, Owen-Smith J. Network dynamics and field evolution: the growth of interorganizational collaboration in the life sciences. *Am J Sociol* (2005) 110(4):1132–205. doi:10.1086/421508

46. Winecoff WK. "The persistent myth of lost hegemony," revisited: structural power as a complex network phenomenon. *Eur J Int Relat* (2020) 26(1_Suppl. l):209–52. doi:10.1177/1354066120952876

47. Nair A, Vidal JM. Supply network topology and robustness against Disruptions–An investigation using multi-agent model. *Int J Prod Res* (2011) 49(5):1391–404. doi:10.1080/00207543.2010.518744

48. Xu J, Zhang B, Lu J, Yu Y, Chen H, Zhou J. Agri-food supply chain concentration and financial growth: configuration and congruence approaches. *Int J Phys Distrib Logist Manag* (2024) 54(9):936–66. doi:10.1108/IJPDLM-01-2024-0038

49. Chen K, Guan J. Mapping the innovation production process from accumulative advantage to economic outcomes: a path modeling approach. *Technovation* (2011) 31(7):336–46. doi:10.1016/j.technovation.2011.03.004

50. Santoleri P, Russo E. Spurring subsidy entrepreneurs. *Res Pol* (2025) 54(1):105128. doi:10.1016/j.respol.2024.105128

51. Carlson JM, Doyle J. Complexity and robustness. *Proc Natl Acad Sci* (2002) 99(Suppl. 1):2538–45. doi:10.1073/pnas.012582499

52. Kitano H. Biological robustness. *Nat Rev Genet* (2004) 5(11):826–37. doi:10.1038/nrg1471

53. Vedres B, Stark D. Structural folds: generative disruption in overlapping groups. *Am J Sociol* (2010) 115(4):1150–90. doi:10.1086/649497

54. Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complex networks. *Nature* (2000) 406(6794):378–82. doi:10.1038/35019019

55. Chen F, Jiang G, Gu M. Supply chain digitization and energy resilience: evidence from China. *Energy Econ* (2025) 144:108420. doi:10.1016/j.eneco.2025.108420

56. Lamin A, Dunlap D. Complex technological capabilities in emerging economy firms: the role of organizational relationships. *J Int Manag* (2011) 17(3):211–28. doi:10.1016/j.intman.2011.05.004

57. Simon HA. The architecture of complexity. *Proc Am Philos Soc* (1962) 106(6):467–482. Available online at: https://www.jstor.org/stable/985254 (Accessed March 31, 2025)

58. Cooke P, Uranga MG, Etxebarria G. Regional innovation systems: institutional and organizational dimensions. *Res Pol* (1997) 26(4-5):475–91. doi:10.1016/S0048-7333(97)00025-5

59. Blanco V, Luthe T, Bruley E, Grêt-Regamey A. Aligning social networks and co-designed visions to foster systemic innovation in the alps. *Reg Environ Change* (2023) 23(3):102. doi:10.1007/s10113-023-02099-y

60. Martin R, Sunley P. On the notion of regional economic resilience: conceptualization and explanation. *J Econ Geogr* (2015) 15(1):1–42. doi:10.1093/jeg/lbu015

61. Rodríguez-Pose A. Do institutions matter for regional development? *Reg Stud* (2013) 47(7):1034–47. doi:10.1080/00343404.2012.748978

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 10 September 2025
DOI 10.3389/fphy.2025.1652953

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Dezhou Kong,
The University of Nottingham Ningbo, China
Haitao Guo,
China University of Petroleum, Beijing, China

*CORRESPONDENCE
Zijing Gao,
2364301732@qq.com

RECEIVED 24 June 2025
ACCEPTED 11 August 2025
PUBLISHED 10 September 2025

CITATION
Lv M, Gao Z, Jiang H, Gao Y and Gao Y (2025) Benefit analysis of regional energy interconnection using system dynamics method targeting carbon neutrality. *Front. Phys.* 13:1652953.
doi: 10.3389/fphy.2025.1652953

COPYRIGHT
© 2025 Lv, Gao, Jiang, Gao and Gao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Benefit analysis of regional energy interconnection using system dynamics method targeting carbon neutrality

Ming Lv[1], Zijing Gao[1]*, Han Jiang[2], Yi Gao[2] and Yuan Gao[3]

[1]College of Biochemical Engineering of Beijing Union University, Beijing, China, [2]Global Energy Interconnection Group Co., Ltd., Beijing, China, [3]State Grid Economic and Technological Research Institute Co Ltd., Beijing, China

The construction of the global energy internet is a significant way to develop clean energy and the key to achieving the goal of carbon neutrality. This study aims to the feasibility of trans-national and trans-continental development of energy interconnection. The trend of regional power grid interconnection is accelerating, and the influence of energy interconnection on global economic and social development is increasing. The prior literature is absent on technical issues such as energy transmission, storage, sharing, and system issues such as power trading, overall architecture, and data collection. This paper presents the contribution of an utterly systematic model and method for evaluating the comprehensive benefits of energy interconnection and analyzes the comprehensive benefits of energy interconnection development from multiple dimensions in economics, society, resources, and links. This paper takes the economic and clean energy development system in Southeast Asia as an example to make the system simulation, designs scenarios of quick, steady and slow carbon peaking & neutrality, and evaluates the comprehensive benefits under the three economic development modes using the evaluation model. The result presented that the development of energy interconnection has an apparent positive effect on economic growth, eliminating the people lacking electricity, poverty, and other social problems, improving resource utilization efficiency, and the ecological environment. The method proposed in this paper for evaluating the benefits of solving energy interconnection issues is practical and feasible and offers adaptable policy insights for regions at different development.

KEYWORDS

regional energy interconnection, carbon neutrality, comprehensive benefits, system dynamics, scenario analysis

1 Introduction

Global energy development is facing severe challenges such as resource shortage, environmental pollution, climate change and unbalanced development, which bring a serious threat to the survival and development for human beings. The Paris Agreement provides a framework for the global response to climate warming after 2020. 178 parties have been implementing carbon emission reductions to promote carbon neutrality and carbon peaks.

In order to achieve the goal of carbon neutrality, the fundamental solution is to accelerate the development of clean energy and the promotion of energy transformation. Building a global energy interconnection is a newly clean and green way to meet the global electricity demand, a global solution to solve the severe challenges of energy development, and a new blueprint for the green and low-carbon development of the world energy. Global energy interconnection is a globally connective, strong and smart grid based on ultra-high voltage backbone network frame, and a basic platform of clean energy for globally large-scale development, transmission, and use. The essence of global energy interconnection is to build up an energy community and human destiny community, which gets strongly supports from United Nations and other relevant parties.

However, in the literature [1], the implementation of the global energy interconnection has faced some resistance and has been questioned in several aspects such as ultra-high voltage long-distance transmission technology, geopolitical resistance and economic and environmental feasibility, which need to be studied systematically and comprehensively. The investment of regional energy interconnection construction is huge. In the short term, the investments are difficult to recoup, but in the middle and long term, regional energy interconnection construction is good to eliminate the population lacking electricity, promote the regional economy and employment, improve the efficiency of energy utilization and reduce the harmful gas emissions. At present, there is no specific and explicit study of the construction effects of energy interconnection on economy, resource, environment and society, and there is also not a set of scientific and systematic model and comprehensive benefit evaluation system of energy interconnection development to make comprehensive study in terms of cross system science, information science and economics.

There are great advantages and feasibility of the System Dynamics theory to analyse the comprehensive benefits of trans-national and trans-continent clean energy transmission channels. First, the System Dynamics theory solves the problem of multi-factor interaction. The trans-national and trans-continent clean energy transmission channels involves subsystems such as economics, environment, resource utilization, scientific and technological innovation, people's livelihood and social problems, and the variables in each subsystem are interaction with complex relationships. Based on this situation, the System Dynamics theory is a structure model showing the relationship between the variables, making the decision and simulating the state transition process of the system based on state variables as the center and speed variables as the drive. In addition, the System Dynamics theory deals with the qualitative information existed in the analysis. Compared with the traditional analysis with no quantitative research for the problem such as social influence and concepts, the System Dynamics theory is a kind of system thinking, system analysis, quantitative and qualitative combined method and integrated and reasoning combined method, making up for the limitation of the traditional quantitative research methods to simulate the complex and large system easily, in particular, analyze and test problems which are difficult to quantification, so that to analyze and research the system in the case of data insufficient. Finally, the System Dynamics theory can be used for policy analysis and simulation. As a soft science research problem, the comprehensive benefits of trans-national and trans-continent clean energy transmission channels can't be tested in the laboratory like natural science. The System Dynamics theory as a "lab" of economic and social problems, has the characters of policy experiment and social laboratory, for example, changing policy variable parameters and variable initial values of the model, making the assumption.

The paper builds a comprehensive benefit evaluation model based on the System Dynamics theory for highly nonlinear, high-order, multi-variable and multi-feedback energy interconnection complex systems. First of all, there is a comprehensively analysis about the effects and benefits brought by the regional interconnection construction in terms of economic, social, resource and environmental aspects. Then, the relationship between each factor and the feedback mechanism are analyzed, the weight is determined by combining the analytic hierarchy process, and the comprehensive benefit evaluation model is established. Finally, the typical situation in Southeast Asia is analyzed and verified through the evaluation model, and then the recommended development model is provided. The findings are that the development of energy interconnection plays a positive role in driving economic growth, eliminating social problems such as population lacking electricity and poverty, and improving the efficiency of resource utilization and ecological environment. There should be scientific positioning and systematic planning when choosing the quick carbon peaking & neutrality scenario.

2 Literature review

This section includes the review on energy interconnection and energy system dynamics model.

2.1 Energy interconnection

The development of global energy interconnection is becoming more and more important to the development of global economic and social, which has received attentions increasingly at present. Wang [2] proposed the solutions of realizing the sustainable development of world energy, and analyzed the comprehensive benefits of constructing the global energy Interconnection [3]. put forward a comprehensive benefits assessment method of different new energy generation technologies for global energy Internet development, in order to evaluate the importance of clean energy to the global energy interconnection. Yang et al. [4] divided the development stages of the global energy interconnection and evaluated the indicators of grid interconnection based on analysis of research data and key indicators. According to cluster analysis, there were six development stages of the global energy interconnection. Electric vehicles were one of the most important links between electric power system and transportation system under the power grid development stage. Dai [5] analyzed the development status of electric vehicles and wireless charging technology in China and provided an innovative idea for the integration of electric power system and transportation system under the future unified power grid framework. Zhou et al. [6] systematically studied the energy interconnection from the perspective of business. The four development stages of energy system were decentralized energy system, centralized energy system, distributed energy system and

intelligent interconnected energy system, while energy network was the innovative development mode of energy system. Wei et al. [7] analyzed the overall framework model of urban energy interconnection from the aspects of energy supply, consumption form, energy transport, decision management and public service. Cheng et al. [8] solved the problem of uncertainty condition of boundary in the power grid planning by collecting interconnection information and studied the data collection methods in terms of renewable energy, climate, economy and energy network planning.

The technology problem of energy interconnection is also a hot topic of energy interconnection research. Alexander et al. [9] believed that energy storage system can be used for short-term control (such as power grid frequency control), medium-term balance mechanism (transforming daily production to meet demand), or long-term storage (seasonal transfer) and instead of grid expansion. Smart grid was the infrastructure of the modern grid, Kolokotsa [10] integrated buildings and communities into smart grid and solved the technological problem of the smart grid. The future smart grid must realize the vision of sharing energy interconnection by millions of people in their own homes, offices and factories with others. Wu et al. [11] believed that grid interconnection technology can transform the power grid into an external grid for energy sharing. Sun et al. [12] evaluated the accuracy of the routing function design and transmission stability of the energy network and verified the proposed energy function sector through theoretical analysis and experimental simulation. Hao et al. [13] proposed two distributed interaction algorithms for both islanded and grid-connected modes where the price signal motivates aggregators can adjust power consumption and generation: one is Privacy-preserving gossip algorithm which can avoid disclosing local privacy in islanded interaction, another is Perturbation primal-dual sub-gradient method which can optimize distributed demand response (DDR) considering global predetermined unequal constraints. Kwon et al. [14] built a variety of scenarios including renewable energy capacity, interconnection measures and reinforcement measures, and then, database creation and simulation are performed semi-automatically to optimize and reinforce future Korean power systems using Power System Simulator for Engineering (PSS/E) and Python by a dynamic model based on inverter-based resources.

Global energy interconnection solved the problem of asynchronous between renewable resource and load center allocation through intercontinental grid interconnection, Liu, et al. [15] put forward the decision framework of power generation and transmission under the background of global energy interconnection, analyzed the influence factors of investment decision, and proposed the corresponding suggestions and models. In order to improve the efficiency and flexibility of energy interconnection, Zhong et al. [16] proposed an SDEI framework within the SDN methodology of energy interconnection and built a programmable structural system in the energy router grid by using the SDN method, which improved the energy efficiency and the green energy utilization. Xi et al. [17] proposed an equilibrium model containing each energy subsystem objective and maximum social welfare for improving the operation of the electricity, gas and district heating subsystems of a district or urban area. Yang et al. [18] modeled the planning profit models of electricity network and gas network and a dynamic game model for the joint planning of electricity-gas integrated energy system based on the analysis of the game mechanism of different investors, at last, the iterative search method was used to solve the model to propose a joint planning method of electricity-gas integrated energy system. Peng et al. [19] constructed distributed integrated energy system by a hierarchically partitioned structure of urban integrated energy system based on distributed autonomy and centralized coordination and a hierarchically partitioned master-slave game optimization model based on game theory.

Liu [20] found that transnational, transcontinental and long-distance power transmission is an important form of global energy interconnection. The development trend of the current stage is to build transnational and transcontinental grid interconnection projects and to form regional energy interconnection at first. In the literature Liu [21], the feasibility and necessity of intercontinental power grid interconnection was demonstrated. Super Grid plans of Europe is to build a power transmission, carrying and consuming platform in Europe; The DESERTEC power project launched by Germany is to develop solar energy resources in Northern Africa and to carry out inter-continental grid interconnection research between North Africa and Europe. The Australian research institute Grenatec Research [22] also proposes the idea of pan-Asian energy interconnection to realize the joint absorption of Australian clean energy in Southeast Asia and even in the whole East Asia.

Some scholars are evaluating which way of energy interconnection is better [23], pointed out the renewable energy based hybrid systems is more feasible through comparing submarine cable interconnection and renewable energy based on hybrid system development for 132 islands of Philippine. Pavievi et al. [24] constructed two models to evaluate the cost-effective decarbonisation strategy of cross-sectoral linkages: one is a long term planning multisectoral model, another is a unit commitment and optimal dispatch model covering multiple energy sectors such as power, heating & cooling, transportation, etc. At last, the data from Europe-wide energy systems showed that the transport sector provides the highest flexibility potential in terms of power curtailment, load shedding, congestion in the interconnection. Wang et al. [25] believe that transnational power grid interconnection, as a measure of energy Internet, can meet the global power demand in a clean and green way and promote the UN's concept of "sustainable energy," but it is very important to build a risk assessment index system including 10 indexes such as national relation, public participation, and available transmission capacity because of the complex and variable risks of transnational power grid interconnection. However, relevant studies only assume a specific case of power interconnection project at present, and there is still lack of systematic evaluation model to quantify the benefits of regional energy interconnection.

The construction of regional energy interconnection brings many comprehensive benefits. In terms of social aspects, in the literature Grenatec Research [22], based on the multi-country and multi-sectoral dynamic conditions, the general equilibrium model is computed, the evaluation index system of economic and social benefits is established, and the quantitative analysis of economic and social benefits of each continent is carried out. In the literature Jin et al. [26], an evaluation system of social benefits is constructed based on the pressure-status-response model, and the impacts of energy interconnection construction on social benefits

through clean energy replacement, power grid optimization and power substitution are evaluated by using smart grid construction data of China.

In terms of economic aspects, Xiaoxia et al. [27], Zhang et al. [28] and [29]. identified the factors effecting the transmission cost and the benefit of energy interconnection. Alves et al. [30] modeled the power systems of the islands of Pico and Faial in Azores, the results showed that RES penetration can increase 50 percentage points by 2030 because of using energy plan to interconnect the isolated power systems of islands. Purvins et al. [31] believe that long distance power transmission is an option for reducing carbon footprint in future electricity systems, such as an electricity trading route crossing the interconnected countries: Turkey, Georgia, Azerbaijan and Kazakhstan, the economic benefit may compensate investments between 1598 M€ and 3251 M€ needed for the additional power transmission capacities over a life-cycle period of 40 years. Breyer et al. [32] pointed out that the benefits of a global energy interconnection is popular, energy interconnected globally would be 4% lower than an isolated system if world is divided into 9 major regions and the major regions comprise of 23 regions.

In terms of the environment aspect, in the literature Zong [33], the self-balancing scenario and power interconnection scenario of global power development are set up. By establishing a global energy interconnection environmental benefit index system, the benefits of power interconnection in terms of resource conservation, convention pollutant emission reduction and climate change, etc. are quantitatively evaluated. In terms of relative benefits aspect, in the literature Zhao et al. [34], the costs and benefits of the global energy interconnection construction and the reforming progresses of the China's power industry marketing characteristics are analyzed and evaluated, and the external effects of the development of energy interconnection are analyzed. IEC [35] analyzes impact of energy interconnection on energy supply, the environment, technologies, policies and standards development. Wang et al. [36] established an optimized model to verify the energy trading interconnection in southern China can yield significant energy, economic, environmental and interconnection benefits. Fan et al. [37] proposed a power gird interconnection model using linear programming techniques to minimize overall system costs due to the difference in the countries' power system, and four power grid interconnection programs in Northeast Asia is proved that power grid interconnection can bring environment and economic benefits. The object of energy interconnection benefit in current research is limited to a certain aspect of economy, society and environment, and the research method is relatively simple, which is lack of a comprehensively analysis on the development of regional energy interconnection.

Current research on the assessment of transnational energy interconnection typically focuses on single-dimensional analyses in the economic or environmental domains, such as quantifying the cost-effectiveness of carbon emission reduction, conducting research on technical feasibility, or testing the stability of cross-border power transmission. The research methods employed are relatively simplistic, failing to provide a comprehensive analysis of the development of regional energy interconnection. Some models have not decoupled the mutual feedback mechanisms between social factors like employment growth and improvements in people's livelihoods, which may lead to imbalances in policy design and consequently overlook social benefits. Meanwhile, certain analytical frameworks, due to their rigid setting of parameters related to economic or environmental policies, cannot adapt to the dynamic changes in technology and policies. Therefore, there is an urgent need to construct a comprehensive and dynamic assessment model and methodology that integrates economic, social, and environmental benefits to analyze the complex interactions and nonlinear effects of various factors in the development of regional energy interconnection.

This study constructs a System Dynamics model targeting employment creation, environmental improvement, and energy equity in Southeast Asia. It innovatively proposes a comprehensive benefit evaluation methodology for regional energy interconnection spanning economic, social, resource, and environmental dimensions. The model explicitly elucidates the energy transition mechanism whereby clean energy investments drive employment growth, increase incomes, and enhance environmental quality. Furthermore, it enables dynamic policy simulations—such as carbon taxes and environmental subsidies—capturing their nonlinear effects under varying economic development levels and clean energy adoption scenarios, overcoming the limitations of static policy assumptions in conventional models.

By integrating social equity into the multi-dimensional benefit assessment framework of energy transition, this research advances a coupled theory of economic-social-energy systems. It provides Southeast Asia with sustainable development pathways that simultaneously meet rigid emission constraints and social equity imperatives, thereby resolving the dilemma between decarbonization and poverty alleviation in developing regions.

2.2 Energy system dynamics model

System dynamics (SD) is a methodology and mathematical modeling technique to frame, understand, and discuss complex issues and problems for understand the behavior of complex systems over time. Wang et al. [38] presented a novel approach which integrates object-orientation and systems-dynamics to information modeling of socio-tech systems. In term of the power system aspect, in the literature Vernotte et al. [39], a system dynamics model was developed for simulation of the power system behavior. [40] produced an efficient state prediction method based on system dynamics, utilizing a statistical approach called the M-Estimation to resolve the filtering problem. Besides [41], developed a simulation system for modeling the behavior of the power plant. Moreover, SD can be also used throughout the public and private sector for policy analysis and design. In the literature of Mora et al. [42], the adequacy of using systems dynamics method to complement the knowledge on the designing making support system implementation processes have been proved.

The remarkable characteristic of this method is that it can integrate non-linear causality and feedback among many factors. Therefore, it has obvious advantages in dealing with comprehensive systems affected by many factors. Gu et al. [43] proposed a model by making an organic integration of system dynamics and entropy weight method, a simulation model for the evaluation of the planning scenario is built based on system dynamics, and

then technical economic indicators with a time scale could be yielded. In the literature Wu et al. [44], the cost of low-carbon power technology and transaction expenses of the carbon trading mechanism was introduced into the planning of interconnection while considering the emergence of new factors and their impact on the power system interconnection under the condition of low-carbon economy. Khansari et al. [45] provided a comprehensive system model, which is based on the system dynamics approach, to address the issue of global warming, in terms of households' energy consumption behaviors. Oluic et al. [46] present a system dynamics simulator to illustrate the behavior of various factors related to the three pillars (social, economic, and environmental) and factor relationships associated to a specific problem.

3 Methods

This section includes the proposed method, Energy System Dynamics model, System hypothesis and parameter setting, the construction of regional energy interconnection System Dynamics model, Main equations of System Dynamics, and Evaluation process of regional energy interconnection comprehensive benefit.

3.1 System dynamics modeling theory

Fan and Pan [47], SD is an analytical technology for quantitative simulating study of nonlinear, multiple feedback and complex time-varying system by computer based on feedback control theory proposed by professor J. W. Forrester of the Massachusetts institute of technology (MIT) (see [48] and [49]). This study is to deal with complex system problems in terms of social, economic, ecological and biological, and comprehensively study the large-scale systems with the characters of complex, multi-level, multi-sectoral and non-linear at macro and micro levels.

Broad sense of system includes open loop system and feedback system, and social and economic system belongs to information feedback system, which is influenced by the historical behaviors of the system and gives the consequences of historical behaviors back to the system correspondently to influence the future behaviors.

SD supposes all of the system motions in the world are fluid motions, and represents the structure of the system using causal and system flow diagrams. The causal diagram can clearly express the nonlinear causal relationship within the system. The causal relationship graph takes the feedback loop as its constituent element and the feedback loop are series of closed paths of causes and results.

The flow graph of the system is the basis of the System Dynamics model, which can represent the feedback relationship in the system, and composed of specific symbols to represent the correlation between the variables in the causal relationship loop. The system flow graph consists of three elements: state variable, rate variable and information, and the relationships between variables are defined by formulations. The relationship between state variables and rate variables is a first-order differential equation, and the entire System Dynamics model is a series of nonlinear equations, as shown in Formula 1:

$$\frac{d}{dx}x(t) = f(x,p) \tag{1}$$

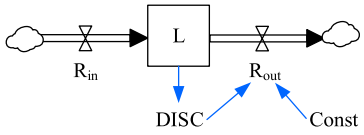


FIGURE 1
Feedback system flow diagram.

Where x is the state variable vector, p is a set of parameters, and f is a nonlinear vector function. The formula is time lag because the vector x and other parameters are functions with the value at the previous moment.

Figure 1 shows the typical flow diagram of a feedback loop system. As shown in the figure, there should be explicit arithmetic expressions of the influx and outflux of substances, energy and information, etc. These expressions form the right term of Equation 2 and the relationships can be shown as integral equal:

$$L(t) = L_0 + \int_0^t \left(\sum R_{in}(t) - R_{out}(t)\right)dt \tag{2}$$

Where $L(t)$ represents the state variable Lat the moment of t, L_0 represents the initial value of L, R_{in} represents the input flow of the state variable and R_{out} represents the output flow of the state variable. The integral equation above shows that the value of the state variable at time t is equal to the initial value of the state variable plus the integral of the net flow changing over time [0, t].

3.2 Applicability analysis

Hou [50] analyzed the regional energy interconnection, which involves various factors of economic and social development, and its benefit evaluation is a complex system problem of highly non-linear, high-order, multi-variable and multi-feedback. The System Dynamics can deal with the problem well and explore the problem at both macro and micro levels. The main expresses are:

Solving the problem of multi-factor interaction - Fan and Pan [47], regional energy interconnection involves the subsystems of economic, environment, resource utilization, scientific and technological innovation, the people's livelihood and society. There are interactions between various factors in the subsystems and the relationships between factors in the subsystems are complex. The structural model of System Dynamics takes state variables as the center and speed variables as the drive, clearly shows the effects and relationships between the factors and applies decision to simulate the state transition process of the system.

Qualitative information in the process and analysis - [51] explored the advantages of System Dynamics. In the traditional analysis, quantitative research of social influence and concept cannot be implemented, but System Dynamics makes up for the limitations of the traditional quantitative research method through system thought, system analysis, quantitative and qualitative combined, integration and inference combined. It is easy for the System Dynamics to simulate the complex and large system, especially involving the problems which cannot be quantified easily into the

system structure to do analysis and test. Besides, the system analysis and research can be implemented in the case of insufficient data.

Policy analysis and simulation - Zuo and Chen [52] discussed comprehensive benefit of regional energy interconnection is a soft science research problem, which cannot be tested in the laboratory like natural science, but System Dynamics as the "laboratory" of economic and social problems, can change the policy variable parameters of the model and the initial value of variables, to carry out the policy environment condition assumption, policy planning, scenario analysis, sensitivity analysis and policy comparison under different conditions. In addition, System Dynamics carries out the dynamic simulation experiment to accurately reflect the system affected by internal and external environment, and the problem solving is straight, vivid, strongly man-machine dialogue functional. It's easily and directly to dialogue with policymakers, which has the characters of policy experiment and social laboratory.

3.3 System hypothesis and parameter setting

3.3.1 Model boundaries

The general steps of System Dynamics model studied by Zhong et al. [16], the boundary of the system should be first determined,The time boundary of the system dynamics model in this study is set as 2005–2050, covering the critical stages of carbon peaking and carbon neutrality in Southeast Asia. The geographical boundary focuses on the 10 Southeast Asian countries. Because the regional energy interconnection is affected by many factors and to highlight the major benefits of the regional energy interconnection, the System Dynamics model can be divided into economic benefit subsystem, environment benefit subsystem, resources benefit subsystem and social benefit subsystem, within them, the four subsystems respectively depict the comprehensive benefits of regional energy interconnection in four aspects of economy, environment, resources, and society. The basic assumptions of the model include:

Assuming that the region's economic development is stable without major social fluctuations, severe economic crisis and other emergencies, multiple regressions can be carried out through the situations of recent 15 years to predict the comprehensive benefits of regional energy interconnection.

Assuming that local governments, power departments and enterprise managers are all "rational brokers" and all of them follow the general laws of economic market development.

Assuming that the government will impose environmental tax and carbon tax in appropriate years in order to protect the local environment and tackle the climate change.

3.3.2 Parameter setting

The basic parameters of the model include: The data about generation capacity of clean energy comes from Southeast Asia Energy Interconnection Planning Research Report (GEIDCO); The data about current situation and forecast of GDP comes from the ASEAN energy center; The data about population status and prediction comes from the United Nations; Trend of Levelized Cost of Energy (LCOE), detailed information is shown in Figure 2. Data come from the Southeast Asia Energy Interconnection Planning Research Report (GEIDCO).

Detailed data information on other model parameters can be found in Supplementary Appendix 1.

3.4 The construction of regional energy interconnection system dynamics model

The benefit analysis of regional energy interconnection is mainly divided into four aspects: economic benefit, social benefit, resource benefit and environmental benefit. The detailed feedback relationships of the four subsystems are shown in the Supplementary Appendix.

3.4.1 Economic benefit subsystem

The construction and development of regional energy interconnection first generates economic benefits and has a strong relationship with local GDP and infrastructure investment. The main feedback relationships in the economic subsystem include:

1. Economic efficiency index → + regional GDP annual growth rate → + regional GDP total amount → + electric power growth GDP → + economic benefit index
2. Carbon dioxide emission quota → + benefits of carbon emission right→ + total profit → + economic benefit index
3. The proportion of clean energy in power infrastructure investment → + the proportion of clean energy generation in total power generation → + clean energy generation → + unit power supply cost of clean energy → + economic benefit index
4. The proportion of clean energy in power infrastructure investment → + the proportion of clean energy generation in total power generation → + the investment of power infrastructure → + the amount of clean energy investment → + unit power supply cost of clean energy → + economic benefit index

3.4.2 Environmental benefit subsystem

Regional energy interconnection development plays an important role in reducing pollutant emissions and protecting the local ecological environment.

The environmental benefit subsystem reflects the overall situation of energy interconnection development in reducing regional pollutant emissions and measures the improvement of regional environment. The main feedback relations in the system include:

1. The proportion of clean energy generation in total power generation → + clean energy generation → + carbon dioxide emissions reduction from alternative clean energy generation → + total carbon dioxide emission reduction → + environmental benefit index
2. The proportion of clean energy generation in total power generation → + clean energy generation → + sulfur dioxide emissions reduction from alternative clean energy generation → + environmental benefit index
3. The proportion of clean energy generation in total power generation → + clean energy generation → + nitrogen oxides

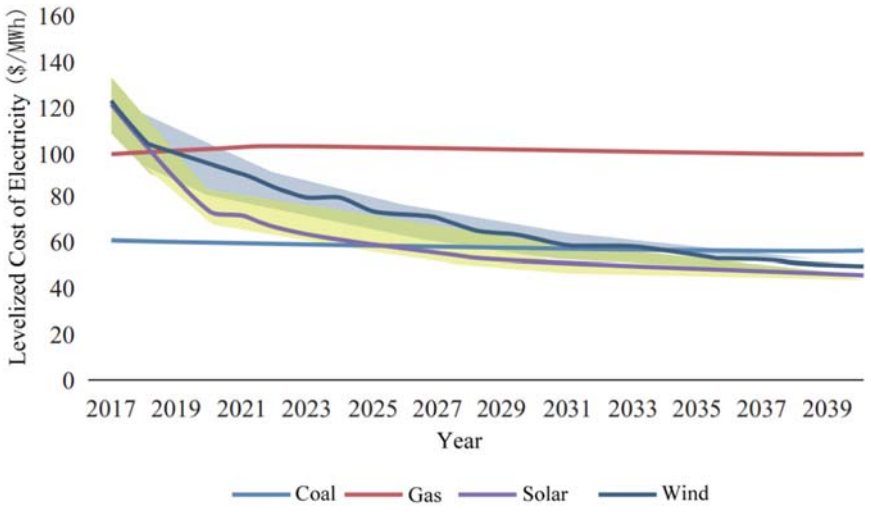


FIGURE 2
Trend of LCOE.

emissions reduction from alternative clean energy generation →+ environmental benefit index

4. The proportion of clean energy generation in total power generation → + clean energy generation → + solid waste emissions reduction from alternative clean energy generation → + environmental benefit index

3.4.3 Resource benefit subsystem

The development of energy interconnection has a significant effect on the regional resource utilization efficiency. By developing clean energy, reducing local fossil energy consumption and water consumption, the local resource utilization efficiency is improves. The main feedback relationships in the resource benefit subsystem include:

1. The primary energy transformation efficiency of clean energy → + primary energy transformation efficiency → + resource benefit index
2. Electric vehicle ownership → + total mileage of electric vehicles → + fuel saving by electric vehicles → + fossil energy consumption reduction by electric vehicles → + resource benefit index
3. The proportion of clean energy generation in total power generation →- the proportion of coal power generation in total power generation → + total resource utilization efficiency → + resource benefit index
4. The increasing amount of electricity consumption in the whole society → + newly power generation → + total power generation → + clean power generation → + fossil energy consumption reduction → + resource benefit index

3.4.4 Social benefit subsystem

The construction of power infrastructure requires a large number of labor forces and technology talents. At the same time, the development of power grid plays an important role in promoting local employment, poverty reduction and resident's livelihood, besides that, it also eliminates the population lacking electricity and improves the living and medical level of the local resident.

The social benefit subsystem is mainly affected by the number of newly employment, the number of people lacking electricity and the proportion of science and research inputs in the total profits. The main feedback relations in the system include:

1. Population growth rate → + regional population → + per capita GDP → + newly employment → + social benefit index
2. The proportion of clean energy in power infrastructure investment → + the proportion of clean energy generation in total power generation → + the newly installed capacity → + the newly employment number → + social benefit index
3. Government levy coefficient → + power public welfare fund → + population lacking electricity → + social benefit index
4. Network loss rate → - total power generation → + the newly installed capacity → + the newly employment number → + social benefit index

As an index to evaluate the comprehensive benefits of establishing the regional energy interconnection, the comprehensive benefit index of regional energy interconnection is affected by economic benefit index, environmental benefit index, resource benefit index and social benefit index and the four benefit indexes are obtained by weighted average. Finally, the System Dynamics causal model can be constructed, and Figure 3 showed a simplified representation of the above relationships from four subsystems.

3.5 Main equations of system dynamics

3.5.1 Economic subsystem

The increasing of electricity consumption GDP_e can be calculated by the whole society electricity consumption

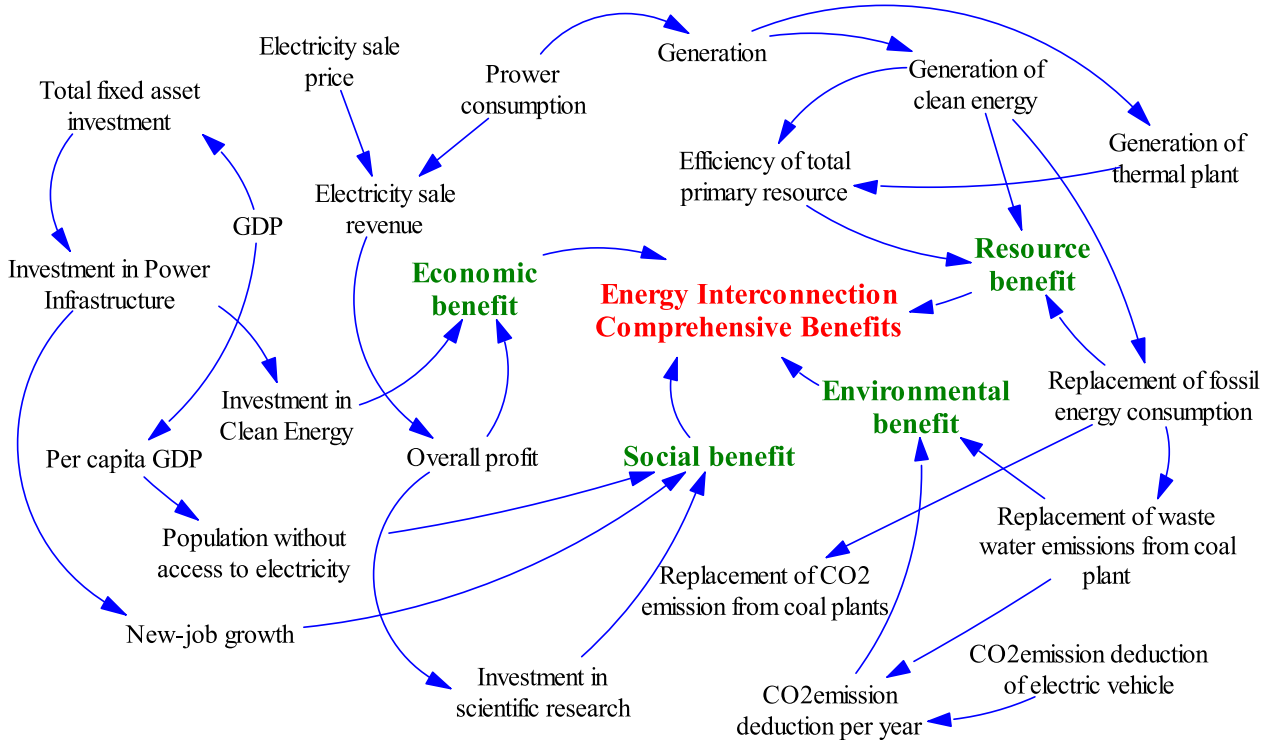


FIGURE 3
System dynamics causal model.

increasement, total electricity consumption and total GDP, as shown in Formula 3:

$$GDP_e = C_{new} \times \frac{GDP}{C_{all}} \tag{3}$$

Among them, GDP_e is the growth GDP of electricity consumption, C_{new} is the increase of social electricity consumption, C_{all} is the total electricity consumption and GDP is the regional total GDP.

Benefits of carbon emission right B_c is determined by the unit price of carbon tax, carbon dioxide emissions and carbon dioxide emission quota. The Formula 4 is as follows:

$$B_c = tax_c \times \left(CO_{2-Q} - CO_{2-coal} - CO_{2-gas}\right) \tag{4}$$

Among them, B_c is benefit of carbon emission right, tax_c is unit price of carbon tax, $CO_{2\text{-}quota}$ is carbon emission quota, $CO_{2\text{-}coal}$ and $CO_{2\text{-}gas}$ are carbon dioxide emissions of coal and gas respectively.

3.5.2 Environment subsystem

Carbon dioxide emissions reduction can be calculated from clean energy generation, as shown in Formula 5:

$$CO_e = G_{cln} \times co_{coal} \tag{5}$$

Among them, G_{cln} is the generation of clean energy and co_{coal} is the carbon dioxide emission per unit of electricity, which is equal to 8.7×10^{-4} tons/kWh, respectively listed by Yao et al. [53].

The reduction of sulfur dioxide emissions is mainly determined by the alterative clean energy generation. The calculation Formula 6 can be expressed as:

$$SO_2 = G_{cln} \times so_{coal} \tag{6}$$

Among them, G_{cln} is the generation of clean energy and so_{coal} is the carbon dioxide emission per unit of electricity, which is equal to 9.0×10^{-6} tons/kWh, respectively listed by Yao et al [53].

3.5.3 Resource subsystem

The primary energy conversion efficiency T_r can be calculated by the proportion of clean energy generation and the proportion of fossil energy generation, as shown in Formula 7:

$$T_r = T_{cln} \times p_c + T_{th} \times p_{th} \tag{7}$$

Among them, P_c is the proportion of clean energy generation and P_{th} is the proportion of fossil energy generation. P_{th} can be further subdivided into two sub-items: the proportion of coal-fired power generation and the proportion of gas-fired power generation, i.e., $P_{th} = P_{coal} + P_{gas}$. The relevant data are all obtained from the Data & Statistics database under the International Energy Agency (IEA). The primary energy conversion efficiency of clean energy T_{cln} is 100%. The primary energy conversion efficiency of coal and natural gas are 32~42% and 32~38% respectively listed by Tan et al. [54].

Fuel saving of electric vehicles O_{EV} can be calculated by the unit fuel consumption of vehicles and the annual mileage of electric vehicles, as shown in Formulas 8, 9:

$$m_{EV} = m_{average} \times n_{EV} \tag{8}$$

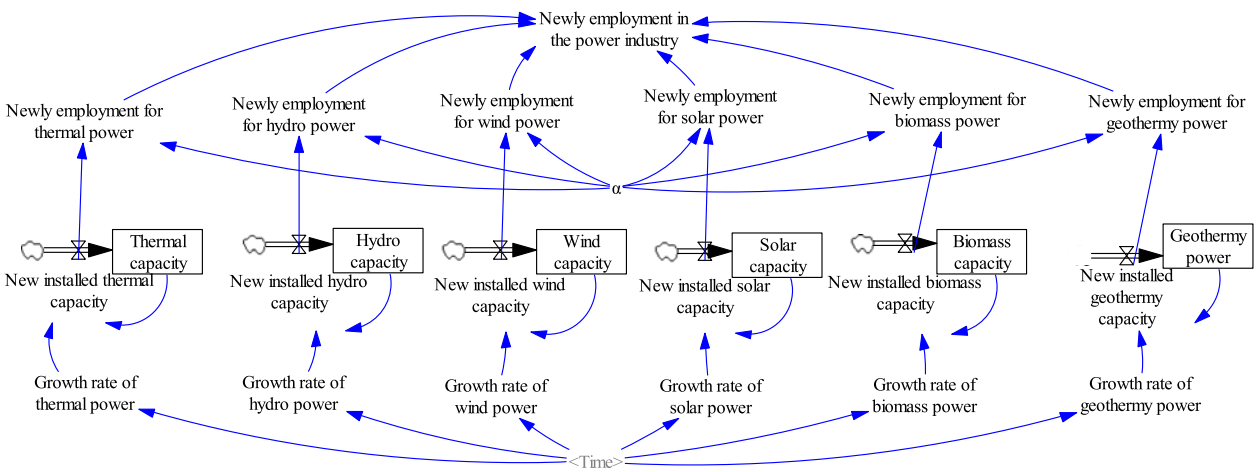


FIGURE 4
Calculation model of new employment.

$$O_{EV} = m_{EV} \times O_{km} \tag{9}$$

Among them, m_{EV} is the annual total mileage of electric vehicles, $m_{average}$ is the annual average mileage of vehicles, n_{EV} is electric vehicle ownership, O_{EV} is the fuel saving of electric vehicles, and O_{km} is the fuel consumption per km of vehicles.

3.5.4 Social subsystem

Research investment is calculated by the total income of the power industry and the research input factors, as shown in Formulas 10, 11:

$$RE_s = PRO_{elec} \times P_{re} \tag{10}$$

$$P_{re} = 0.02 \times p_c \tag{11}$$

Among them, P_{re} is the scientific research input factor, p_c is the proportion of clean energy generation, and PRO_{elec} is the total income of the power industry.

The newly employment in the power industry is determined by the annual increase of various types of installed power facilities and the production, construction, operation, maintenance and fuel supply factors of each type of installed power facilities, as shown in Figure 4. Further considering the technological progress factor and introducing attenuation factor, the calculation Formula 12 of newly employment in the power industry is as follows:

$$J_e = \sum C_x \cdot J_x \cdot \alpha_x \tag{12}$$

Among them, C_x is the new installed capacity of a certain energy type, J_x is the people number factor required by the production, construction, operation and fuel supply of the corresponding energy type, and α is the corresponding attenuation factor.

$$\alpha_x = \alpha_{x\text{-}tech} \cdot \alpha_{x\text{-}lab} \quad \alpha_{x\text{-}lab} = 1 - \beta \tag{13}$$

The above parameters are specifically applied in Formula 13 to calculate the newly employment in the power industry. The attenuation factor α_x integrates the dual impacts of technology and the labor market. $\alpha_{x\text{-}tech}$ measures the degree to which automation and AI technologies lead to the replacement of jobs in the power industry, reflecting the impact of technology on labor demand. $\alpha_{x\text{-}lab}$ is related to the labor market, and the unemployment rate is incorporated into the adjustment through the coefficient β. The higher the unemployment rate, the smaller $\alpha_{x\text{-}lab}$ is. The attenuation factor as a whole decreases as the unemployment rate rises, reflecting that an abundant labor supply will correspondingly reduce the labor demand in the power industry. The above data are from the Asia-Pacific Employment and Social Outlook 2024 released by the International Labour Organization (ILO).

3.6 Evaluation process of regional energy interconnection comprehensive benefit

This study uses regression analysis and analytic hierarchy process to quantify the comprehensive benefits of regional energy interconnection in terms of economy, society, environment and resources. The overall research process is shown in Figure 5.

The equations between different factors can be obtained by regression analysis of historical data, and the relational expression can be calculated by the regression analysis tool. If Multiple R is greater than 0.95 and the error P value is less than 0.05, it can be considered that there is a linear relationship between the indicators and the relational expression can be determined.

The paper uses the maximum and minimum normalization method to obtain the normalization results of the related indexes of each benefit index. The principle of the method is to carry out linear transformation of the original data in order to make the final result falling within the range [0,1]. The transformation formula is as follows:

If the index x is positive related to the benefit index, this is calculated using Formula 14,

$$X_n = \frac{(x_n - \min(x))}{(\max(x) - \min(x))} \tag{14}$$

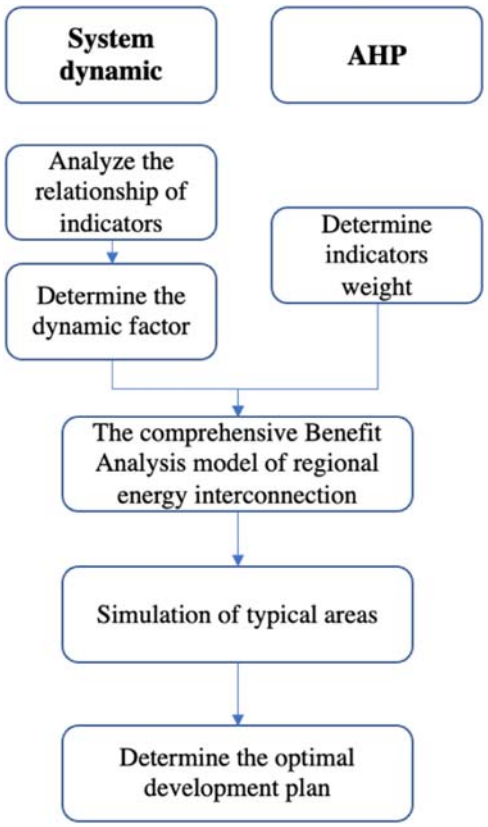


FIGURE 5
Research flow diagram.

If the index x is positive or negative relative to the benefit index, this is calculated using Formula 15,

$$X_n = \frac{(\max(x) - x_n)}{(\max(x) - \min(x))} \tag{15}$$

Among them, X_n is the normalized result with the relative benefit index, and x_n is the index value in the year n, $\min(x)$ and $\max(x)$ are the minimum and maximum values among all of the indicators respectively defined by He [55].

The weight of each factor is determined by the analytic hierarchy process [56]. By means of the score results of relevant experts, the weight of each benefit index is calculated by comparing the relations of importance among various factors, then further obtains the calculation results of each benefit index.

4 Results

4.1 System authenticity inspection

The overall model of the study is based on the Vensim software, and the input variables are related parameters, which are established in terms of equations within the software. To meet the requirements for research reproducibility, some of the data involved in the simulation process have been collated, and the specific contents are shown in Supplementary Appendix 2, 3. Brighthub Engineering and [57], in order to verify the effectiveness of the regional energy Interconnection comprehensive benefit analysis model, and to better use the System Dynamics model for predicting the changing trend of the regional energy Interconnection comprehensive benefit, the study inspects the authenticity of variables simulated by the system from 2005 to 2015, and shrinks the related error of simulation data fitting with historical data within 5% by modifying and improving various indicators through inspections. To demonstrate the degree of the model's fit to historical data, key quantitative validation indicators are listed in Supplementary Appendix 4 to supplement the model validation content.

4.1.1 Sensitivity

This section conducts a systematic sensitivity analysis to address the uncertainties in model parameters and enhance the credibility of the research. The key parameters, including the clean energy proportion (P_c) and GDP growth rate (g), are selected for analysis, and the sensitivity coefficient S is used to quantify the impact of parameter changes on the comprehensive benefit.

4.1.1.1 Calculation of sensitivity coefficients

The sensitivity coefficient S is calculated using the Formula 16:

$$S = \frac{\text{Change Rate of Comprehensive Benefit}}{\text{Change Rate of Parameter}} \tag{16}$$

When the clean energy proportion changes from 30% to 50%, the comprehensive benefit increases by 15.2%. The sensitivity coefficient is calculated as:

$$S_{P_c} = \frac{15.2\%}{20\%} = 0.76$$

When the GDP growth rate changes from 8.0% to 5.0%, the comprehensive benefit increases by −9.7%. The sensitivity coefficient is calculated as:

$$S_g = \frac{-9.7\%}{3\%} = -0.323$$

4.1.1.2 Threshold and sensitivity classification

Following the criterion proposed, a sensitivity coefficient $|S| > 0.3$ is defined as a strong sensitivity.

4.1.1.3 Analysis of results

With a sensitivity coefficient of 0.76 (>0.3), the clean energy proportion has a significant impact on the comprehensive benefit. Adjusting the clean energy proportion can effectively influence the comprehensive benefit, which is crucial for policy-making aimed at improving comprehensive benefits. However, attention should also be paid to potential side effects, such as increased costs.

A sensitivity coefficient of 0.323 (>0.3) indicates that the GDP growth rate is also a critical factor affecting the comprehensive benefit. Rapid economic growth can drive up the comprehensive benefit.

In summary, this sensitivity analysis identifies the strong sensitivity of the clean energy proportion and GDP growth rate to the comprehensive benefit. These findings highlight the need for precise setting of these key parameters in the model and provide a basis for addressing parameter uncertainties. By clarifying the impact of parameter changes on the results, the credibility and robustness of the model are enhanced.

4.2 Multi-scenario index analysis of energy interconnection in Southeast Asia

4.2.1 Scenario assumptions

This study constructs three scenarios: Quick Carbon Peaking, Slow Carbon Peaking, and Baseline Scenario. The core parameters are closely anchored in authoritative regional plans, international institutional forecasts, and measured data to ensure the scientific rigor and policy relevance of the scenario settings.

In the Quick Carbon Peaking scenario, the clean energy structure is strictly aligned with the Global Clean Energy Industry Development Report (2024–2025) and calibrated with regional measured data.

In the Slow Carbon Peaking scenario, in accordance with the World Energy Outlook, the "Baseload Minimum Security Threshold" is strictly implemented, which is consistent with the current demand for stability in industrial electricity use in Southeast Asia.

For parameters related to GDP growth rate and electricity demand, international authoritative institutional forecasts are cited and transformed through the "population-urbanization-electricity elasticity" correlation model.

Based on the development characteristics of energy Interconnection and the forecast of future development policies in Southeast Asia, the development scenarios are proposed for simulation.

1. The first scenario is developing clean energy at high speed, and achieving the goal of carbon peaking and carbon neutrality quickly (or quick carbon peaking & neutrality scenario).

Gao et al. [58] proposed that since clean energy plays a vital role in the environmentally friendly development, there should be positive measures taken to promote the development of the clean energy, and put it into quick carbon peaking & neutrality scenario in this case.

2. The second scenario is developing the clean energy and thermal power in a balanced way, and achieving the goal of carbon peaking and carbon neutrality steadily (or steady carbon peaking & neutrality scenario).

In consideration of the current situation, the government and relevant departments shall take corresponding measures to promote the development of clean energy, making the clean energy generation and thermal power generation in an overall balanced state, and achieve carbon peaking and carbon neutrality steadily.

3. The third scenario is developing the thermal energy, and achieving the goal of carbon peaking and carbon neutrality slowly (or slow carbon peaking & neutrality scenario).

Clean energy will be in a disordered development state without relevant policies and measures., there will be a high thermal power situation generally based on the current development situation. Carbon peaking and carbon neutrality would be achieved slowly in this scenario.

The generation capacities of clean energy in the three scenarios are shown in Table 1; GDP, population and LCOE are shown in Table 2, Table 3 and Figure 2.

4.2.2 Scenario simulation

Through analogously simulating the system and comparing and analyzing the key indexes of each subsystem:

4.2.2.1 Economic benefit

4.2.2.1.1 The GDP growth of electricity power. Vigorous development of clean energy will create a complete industrial chain covering materials, high-end equipment and intelligent manufacturing, while promoting the development of emerging industries such as information and communication systems. By 2050, the GDP growth of electricity power in the three scenarios is expected to be $1.4 trillion, $1.1 trillion and $0.8 trillion respectively.

4.2.2.1.2 Benefits of carbon emission right. By the year 2050, the carbon emission rights benefits of the three scenarios are expected to reach to $11.3 billion, $6.9 billion and $1.8 billion respectively in Southeast Asia.

4.2.2.1.3 Generation cost. The regional interconnection of countries with different resource endowments will promote the large-scale development and utilization of clean energy with lower marginal costs, which can reduce the price of electricity across the region effectively, and further reduce the cost of social and economic development. By 2050, the cost of leveling cost per kW hour of the electricity under the quick carbon peaking & neutrality scenario will be about 2 cents lower than it was in 2015.

4.2.2.2 Environmental benefit

4.2.2.2.1 The reduction of carbon dioxide emission. By the year 2050, the reduction quantities of carbon dioxide in the three scenarios are expected to reach to 1.7 billion tones, 1.2 billion tones and 0.7 billion tons respectively in Southeast Asia.

4.2.2.2.2 The reduction of sulfur dioxide emission. By the year 2050, the reduction quantities of sulfur dioxide in the three scenarios are expected to reach to 1.8 million tones, 1.2 million tones and 0.7 million tons respectively in Southeast Asia.

4.2.2.3 Resources benefit

4.2.2.3.1 The energy primary conversion efficiency. With the proportion of clean energy generation increasing, the energy primary conversion efficiency in Southeast Asia will also increasing. By 2050, the energy primary conversion efficiency of the three scenarios are expected to reach to 73%, 63% and 51% respectively.

4.2.2.3.2 Fuel saving of electric vehicles. With the promotion of electric vehicles, the number of electric vehicles has increased, and the amount of fuel saved by electric vehicles has significantly increased. It is estimated that the number of electric vehicles in Southeast Asia will reach to 136 million by 2050, which will save 366 million tons of fuel.

4.2.2.4 Social benefit

4.2.2.4.1 Employment opportunities. By the year 2050, the accumulated newly employment in the three scenarios in Southeast Asia will reach to 5 million, 4 million and 2.5 million respectively.

TABLE 1 Power generation mix in Southeast Asia.

Unit(TWh)	Quick carbon peaking & neutrality scenario	Steady carbon peaking & neutrality scenario	Slow carbon peaking & neutrality scenario
Total Power Generation	3,138	3,173	3,098
Thermal Power	1,205	1832	2,326
Hydro Power	790	586	335
Solar Power	359	261	120
Wind Power	320	199	104
Geothermal Power	147	88	58
Nuclear Power	86	76	37
Biomass Power and Else	232	130	117

TABLE 2 GDP forecast in Southeast Asia.

GDP (billion USD)	Year 2015	Year 2035	Year 2050
Cambodia	45.4	163.4	300.4
Laos	32.4	124.7	200.6
Myanmar	496.4	1929.4	4,624
Thailand	930.8	1917.4	2881.7
Vietnamese	462.8	1489.6	2373.8
Brunei	25.5	31.1	36.2
Indonesia	2382.2	6086.9	10113.4
The Philippines	6,210	2230.8	4375.3
Malaysia	676.5	1300.4	1819.3
Singapore	396.1	588.6	754.1

TABLE 3 Population forecast in Southeast Asia.

Population (thousand)	Year 2015	Year 2035	Year 2050
Cambodia	15,520	19,720	22,020
Laos	6,660	8,420	9,160
Myanmar	52,400	60,430	62,360
Thailand	68,660	69,200	65,370
Vietnamese	93,570	108,990	114,630
Brunei	420	510	540
Indonesia	2,58,160	3,04,760	3,21,550
The Philippines	1,01,720	1,32,670	1,51,290
Malaysia	30720	38,380	41,730
Singapore	5,540	6,480	6,570

4.2.2.4.2 Population lacking electricity. Under the scenario of quick carbon peaking & neutrality, the population lacking electricity in Southeast Asia will be basically eliminated by the year 2030. Under the other two scenarios, power supply popularization will be delayed to 2040 and 2045 respectively.

4.2.3 Dynamic action pathway

Between 2030 and 2035, the decline in clean energy costs became a key turning point. The costs of solar and wind power dropped by 30.6% and 28.6% respectively, while coal-fired power costs remained stable. This gap stems from factors such as technological upgrading and large-scale production. At the technical level, an 82% increase in clean energy investment drove its power generation share to rise from 35% to 58%, which in turn reduced energy consumption per unit of GDP and significantly improved efficiency. At the policy level, the implementation of carbon taxes further increased the cost of fossil energy, amplified the competitive advantages of clean energy, and accelerated the accumulation of carbon emission rights revenue, forming a synergy between policy tools and market mechanisms.

In the fast-carbon scenario, clean energy investment drove the development of microgrids, which significantly reduced electricity access costs and accelerated power coverage. The stable-carbon scenario was constrained by the lag in the construction of main grids for thermal power, resulting in slow progress in power supply. In the slow-carbon scenario, due to the lock-in effect of coal-fired power, energy investment was concentrated in traditional fields, leading to a significant decline in electricity access rates. Under different paths, energy structures have different effects on employment promotion. The full-industry-chain layout of clean energy can create more

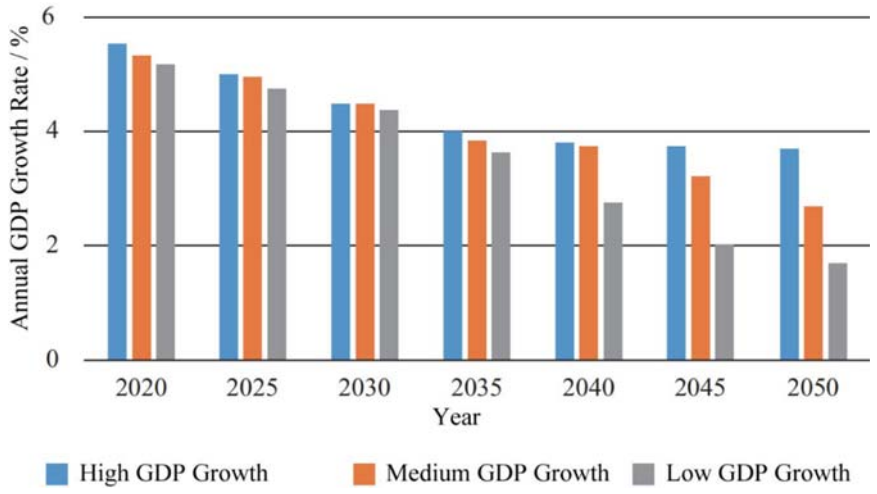


FIGURE 6
Setting of annual growth rate for regional GDP.

synergistic employment opportunities, while the path dependence on traditional energy limits the release of employment potential.

After 2040, environmental benefits improved at an accelerated pace. When the share of clean power generation crossed the critical threshold, the low-carbon attribute of the energy system was enhanced, and the emission reduction effect was significantly improved. Meanwhile, the large-scale popularization of electric vehicles drove the rapid rise in fuel substitution rates in the transportation sector, and advances in battery technology and cost reduction supported this transformation. The synergistic carbon reduction in these two sectors jointly pushed environmental benefits into a stage of leapfrog growth.

4.3 Combination analysis and simulation of energy interconnection development model in Southeast Asia

After simulating the regional energy Interconnection with the single factor, relevant variables can be changed and by combining the annual growth rates of the regional GDP with various development models, different control scenarios will be formed. Through the System Dynamics simulation of the control scenarios and the comparison of various scenarios, the suitable measures to be implementation can be analyzed under various models.

The annual growth rates of regional GDP is set into three different growth models as high, medium and low which is listed by Yu et al. [59], as shown in Figure 6.

Three different annual regional GDP growth rates models and three different regional energy Interconnection scenarios, nine different regional energy Interconnection combination schemes can be generated, as shown in Table 4.

The regulatory political combination plans of regional energy Interconnection scenarios are shown in Table 4. The corresponding control parameters are set for simulating, and the simulation results of each variable under each scheme are obtained by adjusting relevant variables in the models, as shown in Figures 7–11.

Through the analysis of Figure 7, the conclusion can be drawn that under the rapid GDP development situation in the future, quick carbon peaking & neutrality scenario is more suitable for the regional economic development. Improving the proportion of clean energy generation is good to the impacts of clean energy on promoting social and economic system, furthermore, accelerates the regional GDP growth by feedback. Through longitudinal comparison, it can be concluded that under the three scenarios, quick carbon peaking & neutrality scenario has the most obvious benefit differences, indicating that the development of clean energy has a significant promotion effect on GDP.

Through the multi-scenario simulations and analyses of the resources and environmental benefits, the quick carbon peaking & neutrality scenario achieves good environmental and resource benefits under the three GDP growth models. With the GDP growth rate is increasing, the increasing benefits of the quick carbon peaking & neutrality scenario are significant in terms of resources and environments. Therefore, under the high GDP growth model, there will be a better improvement of resources and environments with the priority development of clean energy. If the economic growth rate slows down, the quick carbon peaking & neutrality scenario still has the best environmental benefits. At this time, the steady carbon peaking & neutrality scenario also has certain resources and environmental benefit and reduces certain clean energy investment relatively.

Through the multi-scenario analyses of social benefits, the conclusion can be drawn that the change of social benefit index is not significant and the reactions of three scenarios are not sensitive when the GDP growth rate is increasing, but under the same GDP growth model, the quick carbon peaking & neutrality scenario has the best social benefits, with the index between 0.82 and 0.91, which is significantly higher than the other two scenarios.

By comparing nine different scenarios, the economic benefit index, resource benefit index, environmental benefit index, social

TABLE 4 Different scenario under distinct regulation policy combination schemes.

Mode Scheme Scenario	Quick carbon peaking & neutrality	Steady carbon peaking & neutrality	Slow carbon peaking & neutrality
High GDP Growth	Scheme 11	Scheme 12	Scheme 13
Medium GDP Growth	Scheme 21	Scheme 22	Scheme 23
Low GDP Growth	Scheme 31	Scheme 32	Scheme 33

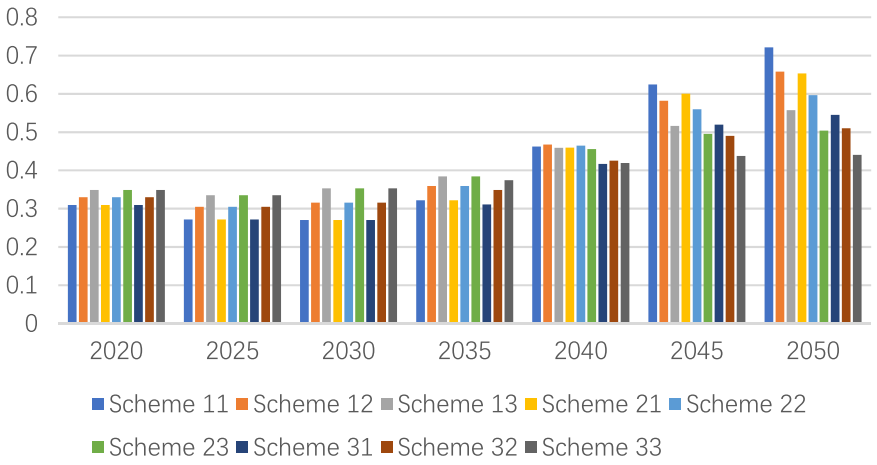


FIGURE 7
Simulation results of economic benefit index under different scenarios.

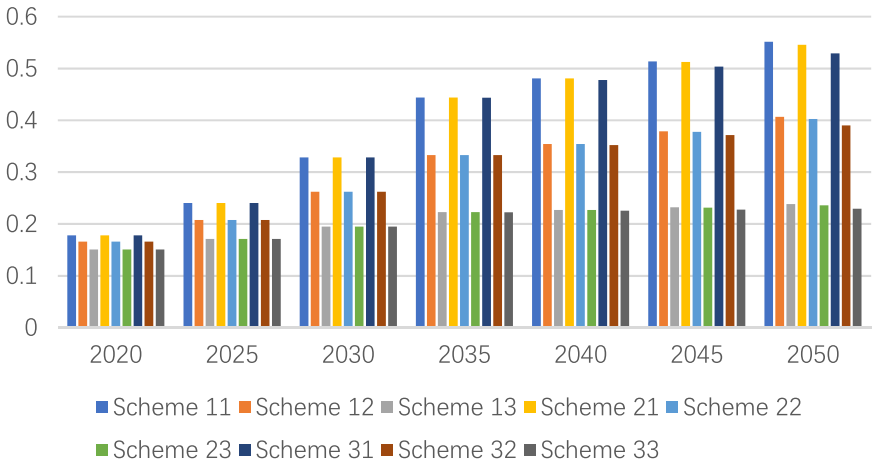


FIGURE 8
Simulation results of resource benefit index under different scenarios.

benefit index and regional energy Interconnection comprehensive benefit index are affected by both economic development level and clean energy situation. Under the scenario of high GDP growth model, the scenario of quick carbon peaking & neutrality scenario has obvious advantages in economic benefit index, resource benefit index, environmental benefit index, social benefit index and regional energy Interconnection comprehensive benefit index, with the highest comprehensive efficiency index of 0.79 until 2050 in scheme 11. Under the scenario of medium speed GDP growth model, the comprehensive benefits of quick carbon peaking & neutrality scenario is still in a dominant position, at the same time, the steady carbon peaking & neutrality scenario is not only good to release

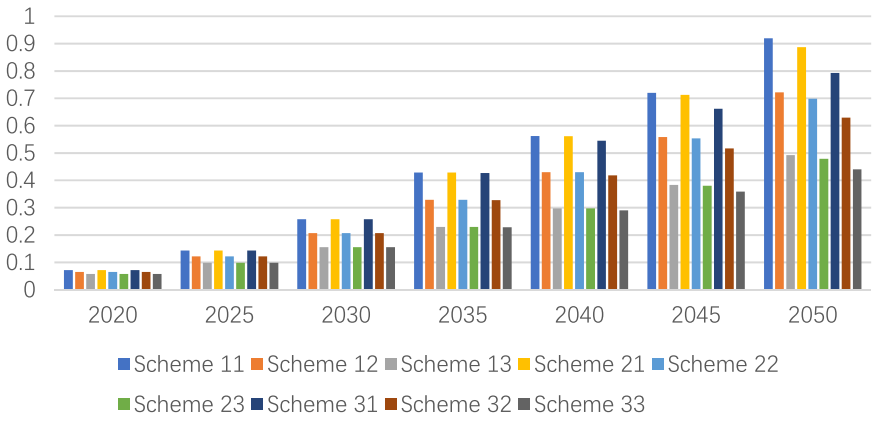


FIGURE 9
Simulation results of environmental benefit index under different scenarios.

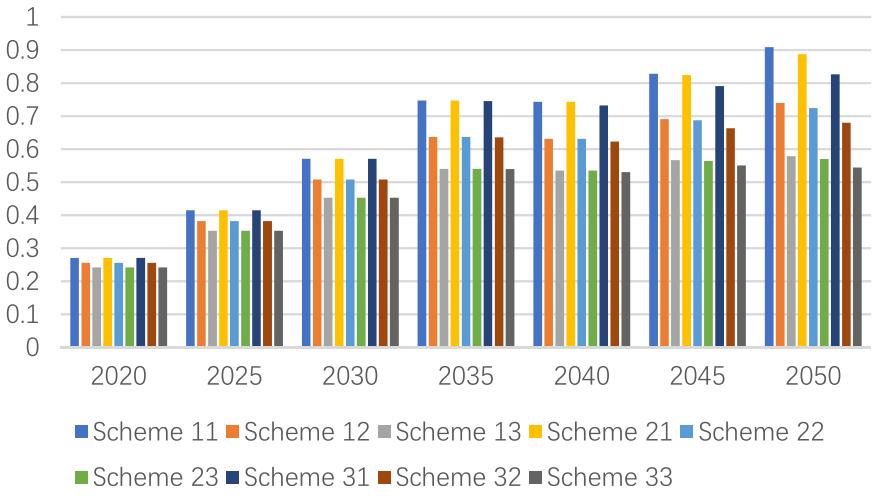


FIGURE 10
Simulation results of social benefit index under different scenarios.

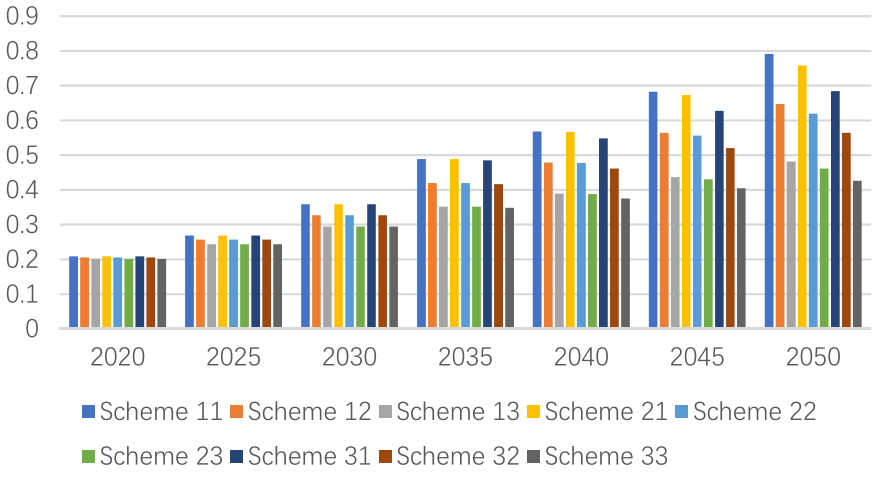


FIGURE 11
Simulation results of comprehensive benefit index under different scenarios.

the investment cost of clean energy construction, but also have good economic, social, resource and environmental benefits, with the efficiency index of 0.59, 0.40, 0.69 and 0.72 respectively. Under the scenario of low GDP growth model, though the comprehensive benefits of the quick carbon peaking & neutrality scenario is the best, its relative advantage is lower than that in the medium and high GDP growth models. Under the circumstance of limited investment, the steady carbon peaking & neutrality scenario can be taken as an alternative.

5 Implications

5.1 Determine the economic, social, resource, and environmental benefits of the energy interconnection

This paper builds a comprehensive benefit evaluation model of the regional energy interconnection, which puts forward the concept of regional energy interconnection comprehensive benefit index and reflects the regional energy interconnection development benefit in the medium and long-term through the economic benefit index, social benefit index, resource benefit index and environment benefit index comprehensively.

This study introduces the data of Southeast Asia into the model for simulation in terms of economic and social development and energy market. The findings are as follows:

1. In the short term, the construction of regional energy interconnection and the development of clean energy replacing traditional energy cannot push the economic growth, but in the medium and long term, it will increase the infrastructure investment, promote the coordinated industrial chain development and increase employment rate. Especially under the background of the carbon tax policy of the future government, the development of energy-saving will improve the profits of relevant energy enterprises.
2. In the medium and long term, the development of regional energy interconnection with clean energy replacing traditional energy as the path will make important contributions to eliminate the population lacking electricity and population living in poverty. It can be found that the population lacking electricity in Southeast Asia will be basically eliminated around 2035 through systematic simulation,
3. Developing clean energy vigorously plays an important role in improving the local energy efficiency and forming a green and sustainable energy utilization pattern. It can be found that the regional energy utilization efficiency can reach to 80% by taking the clean energy replacement as the path around 2040 through systematic simulation.

Studies by the International Renewable Energy Agency (IRENA) and the ASEAN Centre for Energy (ACE) further clarify that the implementation of regional energy interconnection projects, the enforcement of mandatory energy efficiency standards, and the promotion of clean energy substitution policies can significantly improve primary energy conversion efficiency, with the improvement range reaching 75%–85%.

4. The development of clean energy replacing traditional energy has played a positive role in CO2 emission reduction, which reduces 2.4 times of the total amount of CO2 than using traditional energy.

In conclusion, in the short term of 5–10 years, the development of regional energy interconnection with clean energy replacing traditional energy causes the increase of investment and reduces the enterprises' profit, but in the long term of 15–30 years, energy interconnection development is good to eliminate the population lacking electricity, eliminate population living in poverty, improve energy utilization efficiency, promote the economic development quality, increase the regional employment rate and reduce CO2 emissions, etc. This paper makes a quantitative analysis of the comprehensive benefits of the above regional energy interconnection development, and clarifies the influence of the energy interconnection on regional economic and social development through the system dynamics model.

5.2 Evaluation of the model

The comprehensive benefit evaluation model of regional energy interconnection constructed in this paper can evaluate the development level of social economics and energy market in a certain region. Moreover, simulate and analyze the future development path and mode. The model has the characteristics as follow:

1. The model using the method of system dynamics theory, systematically and fully reflects the energy interconnection development benefits from the medium and long term, coordinated and unified the development of energy interconnection with the regional economic, social and environment, determines the interaction mechanism and finds the benefit of energy interconnection on local economics, society, environment and resource in the medium and long term.
2. The model combines quantitative analysis with qualitative analysis and reflects the development of regional energy interconnection objectively through the authenticity test. Meanwhile, due to the uncertainty of the development of the future world, there are different development scenarios of regional GDP development level. The model conducts a comparative analysis of multiple scenarios for different scenarios of regional GDP development, full plays the advantages of system dynamics as "laboratory" in terms of economic and social policy and clarifies the development path of the regional energy interconnection.
3. The model is certainly applicable and universal. In Southeast Asia, due to the development levels of economic and social are balanced and the development foundation of hydropower is good, countries are exploring the typical energy interconnection development mode with clean energy replacing traditional energy. Therefore, through the model, the conclusion can be drawn based on collecting relevant data, simulation analysis and policy analysis. In the future, the model can be used to evaluate the comprehensive benefits of energy

interconnection in other regions, such as Latin America, West Asia and North Africa, etc.

6 Conclusion

This study is based on System Dynamics, constructs the comprehensive benefit analysis model of regional energy interconnection and simulates the system using the example of economic and clean energy development system in Southeast Asia. Besides that, the study designs three carbon peaking & neutrality scenarios, evaluates the comprehensive benefits under the three economic development modes using the evaluation model, and identifies the optimal development mode and path under different scenarios.

The findings are that the development of energy interconnection plays a positive role in driving economic growth, eliminating social problems such as population lacking electricity and poverty, and improving the efficiency of resource utilization and ecological environment. Under the same level of economic development, the quick carbon peaking & neutrality scenario should be chosen, at the same time, regional economic development level plays a positive role in promoting the development of regional energy interconnection. Under the same economic development level, comparing the different carbon peaking & neutrality scenarios, it can be find that the quick carbon peaking & neutrality scenario has a greater effect on the resource and environmental benefit, showing that the development of clean energy is good for the environment, with the GDP growth increasingly, the sensitivity of economic benefits index in quick carbon peaking & neutrality scenario is good, which will increase rapidly. When GDP growth is low, the dominant position of the quick carbon peaking & neutrality scenario declines, the steady and slow carbon peaking & neutrality scenario can be used as an alternative because they can alleviate investment pressure. Therefore, there should be scientific positioning and systematic planning when choosing the quick carbon peaking & neutrality scenario.

Given the differences in development stages and energy demands among Southeast Asian countries, to enhance the relevance and operability of policy recommendations, this section analyzes policies from three dimensions: resource classification, development stage, and model parameters. All recommendations are based on the empirical data and model conclusions of this paper.

For fossil energy-dependent countries such as Indonesia and Vietnam, which have a high proportion of coal-fired power and prominent carbon emission intensity, it is recommended to implement a "coal-fired power capacity replacement mechanism," i.e., for every 1 GW of new clean energy added, 0.8 GW of coal-fired power should be phased out. At the same time, relying on regional advantages, cross-border carbon capture hubs should be established, and CO_2 storage revenues should be incorporated into the regional compensation mechanism to balance the transformation costs.

For countries with high clean energy potential such as Laos and Myanmar, Laos has hydropower potential exceeding 26GW, and Myanmar is rich in photovoltaic resources. These countries are suitable for prioritizing the development of "hydropower-photovoltaic complementary projects" and supporting the China-Laos-Thailand transnational UHV transmission corridors. A "clean energy export tax rebate" policy can be established to reduce electricity costs through cross-border power transmission and encourage cross-border green electricity trade.

For urban high-load countries such as Singapore and Malaysia, Singapore has concentrated electricity load and limited clean energy resources. It is recommended to promote the "distributed photovoltaic + electric vehicle" model, and at the same time implement dynamic carbon taxes on high-energy-consuming fields such as data centers, and improve system flexibility by combining the fuel replacement benefits of electric vehicles.

Due to the differences in economic development levels among countries, the pace of transformation must be gradual to avoid a "one-size-fits-all" approach.

For underdeveloped countries such as Cambodia and Laos, the core contradiction lies in the balance between carbon reduction and poverty alleviation. In the short term (2025–2030), it is advisable to adopt the "stable carbon peak" model (Scheme 22), retain 30% of coal-fired power to ensure employment, and allocate 30% of clean energy investment to rural power grids. In the medium and long term, gradually shift to the "fast carbon peak" model, and absorb green electricity from countries such as Thailand through cross-border interconnection.

For emerging industrial countries such as Vietnam and Indonesia, which face conflicts between industrial electricity demand growth and environmental constraints, it is necessary to implement an "industrial green electricity quota system," requiring high-energy-consuming industries such as electronics and textiles to have a green electricity proportion of no less than 40%. At the same time, a floating carbon tax mechanism should be introduced for steel mills and chemical plants. When GDP growth exceeds 7%, the tax rate should be increased to balance economic growth and emission reduction goals.

The contribution of the study is to propose a completely systematic model and method for evaluating the comprehensive benefits of energy interconnection and conduct a comprehensive benefit analysis of the development of energy interconnection from multiple dimensions in terms of economic, society, resources and links, which plays a guiding role in analyzing the comprehensive benefits of regional energy interconnection construction in the future.

The limitations of the study are that the applicability of the model needs to be further improved, some data are derived from the existing data, and the accuracy needs to be improved. The future study is to tease the influence of the development of energy interconnection on the entire energy industry chain further, build a complex system for the development of energy interconnection, study the characteristics and mechanism of the complex system, and analyze the influence of the development of energy interconnection more systematically and comprehensively.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

ML: Data curation, Resources, Writing – original draft, Funding acquisition, Writing – review and editing, Formal Analysis. ZG: Conceptualization, Software, Writing – original draft, Formal Analysis. HJ: Investigation, Writing – review and editing, Data curation, Formal Analysis. YiG: Writing – original draft, Supervision, Validation, Software. YuG: Formal Analysis, Writing – original draft, Conceptualization, Validation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The Beijing Social Science Fund Project is a provincial and ministerial-level philosophical and social science research project established by the Beijing Municipal Government. This study was supported by the Beijing Social Science Fund Project (22GLC074).

Conflict of interest

Authors HJ and YiG were employed by Global Energy Interconnection Group Co., Ltd.

Author YuG was employed by State Grid Economic and Technological Research Institute Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2025.1652953/full#supplementary-material

References

1. Cao J, Sun J. *Energy interconnection and energy system*. Beijing: China Electric Power Press (2016).

2. Wang Y. Concept and prospect of global energy Interconnection. *China Power* (2016) 49(3):1–6. doi:10.11930/j.issn.1004-9649.2016.03.001.05

3. Yang B, Yao L, Yan S, Li N, Cao Y, Cui H, et al. Comprehensive benefits assessment of different new energy generation technologies for global energy Internet development. *Mater Sci Eng* (2018) 452(3):032036. doi:10.1088/1757-899x/452/3/032036

4. Yang JR, Yang J, Gao CW, Cao Y, Tian W, Ma WQ, et al. The division of developmental stages of the global energy interconnection. In: *Systems and Informatics (ICSAI), 2017 4th international conference on*. IEEE (2017). p. 418–23.

5. Dai B. Prospect of electric vehicle charging station communication mode oriented to energy internet. *Power Inf Commun Technol* (2018) 16(09):8–13.

6. Zhou K, Yang SL, Shao Z. Energy interconnection: the business perspective. *Appl Energy* (2016) 178:212–22. doi:10.1016/j.apenergy.2016.06.052

7. Wei H, Jiang B, Wei XD, Zhu WB, Zhang Y. Research on development strategy of urban power grid based on energy interconnection construction. *DEStech Trans Comput Sci Environ* (2016). doi:10.12783/dtcse/aics2016/8176

8. Cheng L, Ji X, Zhang F, Liang C, He H. Interconnection information applied in the energy interconnection planning: a review and outlook. In: *Energy interconnection and energy System integration (EI2), 2017 IEEE conference on*. IEEE (2017). p. 1–5.

9. Alexander MJ, James P, Richardson N. Energy storage against interconnection as a balancing mechanism for a 100% renewable UK electricity grid. *IET Renew Power Generation* (2015) 9(2):131–41. doi:10.1049/iet-rpg.2014.0042

10. Kolokotsa D. The role of smart grids in the building sector. *Energy and Buildings* (2016) 116:703–8. doi:10.1016/j.enbuild.2015.12.033

11. Wu FF, Varaiya PP, Hui RS. Smart grids with intelligent periphery: an architecture for the energy interconnection. *Engineering* (2015) 1(4):436–46. doi:10.15302/J-ENG-2015111

12. Sun Q, Zhang YB, He HB, Ma DZ, Zhang HG. A novel energy function-based stability evaluation and nonlinear control approach for energy interconnection. *IEEE Trans Smart Grid* (2017) 8(3):1195–210. doi:10.1109/TSG.2015.2497691

13. Hao R, Ai Q, Guan T, Cheng Y, Wei D. Decentralized price incentive energy interaction management for interconnected microgrids. *Electric Power Syst Res* (2019) 172:114–28. doi:10.1016/j.epsr.2019.03.004

14. Kwon HI, Cho YS, Choi SM. A study on optimal power system reinforcement measures following renewable energy expansion. *Energies* (2020) 13:5929. doi:10.3390/en13225929

15. Liu JH, Niu Y, Liu JY, Zai WJ, Zeng PL, Shi HB. Generation and transmission investment decision framework under the global energy interconnection. In: *Power and energy engineering conference (APPEEC), 2016 IEEE PES Asia-Pacific* (2016). p. 2379–84.

16. Zhong WF, Yu R, Xie SL, Zhang Y, Tsang DHK. Software defined networking for flexible and green energy internet. *IEEE Commun Mag* (2016) 54(12):68–75. doi:10.1109/mcom.2016.1600352cm

17. Xi Y, Zeng Q, Chen Z, Lund H, Conejo AJ. A market equilibrium model for electricity, gas and district heating operations. *Energy* (2020) 206:117934. doi:10.1016/j.energy.2020.117934

18. Yang N, Huang Y, Dong B, Xin P, Li H. Research on the joint planning method of electricity-gas integrated energy system based on multi-agent game. *Zhongguo Dianji Gongcheng Xuebao/Proceedings Chin Soc Electr Eng* (2019) 39(22):6521–33. doi:10.13334/j.0258-8013.pcsee.182021

19. Peng LA, Zw A, Wy B, Hlc D, Yy A, Jw A, et al. Hierarchically partitioned coordinated operation of distributed integrated energy system based on a master-slave game - sciencedirect. *Energy* (2020) 214. doi:10.1016/j.energy.2020.119006

20. Liu ZY. *Global energy interconnection*. Beijing: China Electric Power Press (2015).

21. Liu ZY. Research of global clean energy resource and power grid interconnection. *Proc Csee* (2016). doi:10.13334/j.0258-8013.pcsee.160788

22. Grenatec Research. *Pan-Asian energy infrastructure-australia*. Grenatecresearch (2013). Available online at: http://grenatec.com/draft-paei-pan-asian-energy-infrastructure-australia/.

23. Cader C, Bertheau P. Electricity sector planning for the Philippine islands: considering centralized and decentralized supply options. *Appl Energy* (2019) 251:113393. doi:10.1016/j.apenergy.2019.113393

24. Pavievi M, Mangipinto A, Nijs W, Lombardi F, Quoilin S, Jiménez Navarro JP, et al. The potential of sector coupling in future European energy systems: soft linking between the dispa-set and jrc-eu-times models. *Appl Energy* (2020) 267:115100. doi:10.1016/j.apenergy.2020.115100

25. Wang H, Jin Y, Tan X. Study on sustainable development of the transnational power grid interconnection projects under diversified risks based on variable weight theory and bayesian network. *Math Probl Eng* (2020) 2020:1–10. doi:10.1155/2020/5361561

26. Jin YM, Tan X, Jiao BQ. Based on calculated model of global energy interconnection economic and society benefits analysis. *Intell Power* (2018) 46(5).

27. Xiaoxia W, Ding J, Liu J, Tiezhong W. Benefit and cost competitiveness analysis of wind and solar power inter-continent transmission under global energy interconnection mode. *Int Conf Energy Eng Environ Prot (Eeep2016)* (2016). doi:10.1088/1755-1315/52/1/012002

28. Zhang F, Xu Z, Jiao B, Yang J, Feng J. Global energy interconnection based optimal power systems planning — a global perspective 2016–2050. In: *2017 China international electrical and energy conference (CIEEC)* (2017). p. 849–54.

29. Dashash MA, Mahfoudhi R. Energy cost optimization through the implementation of cogeneration and grid interconnection. In: *IEEE power engineering Society general meeting, Montreal, que. 2006* (2006). p. 5.

30. Alves M, Segurado R, Costa M. Increasing the penetration of renewable energy sources in isolated islands through the interconnection of their power systems. The case of pico and faial islands, azores. *Energy* (2019) 182:502–10. doi:10.1016/j.energy.2019.06.081

31. Purvins A, Gerbelova H, Sereno L, Minnebo P. Social welfare impact from enhanced Trans-asian electricity trade. *Energy* (2020) 215:119106. doi:10.1016/j.energy.2020.119106

32. Breyer C, Bogdanov D, Aghahosseini A, Gulagi A, Fasihi M. On the techno-economic benefits of a global energy interconnection. *Econ Energy Environ Pol* (2020) 9(1). doi:10.5547/2160-5890.9.1.cbre

33. Zong HJ. The mechanism and benefit calculation of global energy interconnection social benefit. In: *Fujian forum, humanities and social edition* 8 (2017).

34. Zhao QL, Feng JS, Jin YM. Environmental benefit based on the global energy interconnection. *Glob Energy Interconnection* (2018) 1(5).

35. International Electrotechnical Commission (IEC). Global energy interconnection. IEC White Paper. (2016).

36. Wang H, Su B, Mu H, Li N, Jiang B, Kong X. Optimization of electricity generation and interprovincial trading strategies in southern China. *Energy* (2019) 174(MAY 1):696–707. doi:10.1016/j.energy.2019.03.008

37. Fan J, Wang X, Huang Q, Zhang X, Li Y, Zeng P. Power gird interconnection with hvdc link in northeast Asia considering complementarity of renewable energy and time zone difference. *The J Eng* (2019) 2019:1625–9. doi:10.1049/joe.2018.8851

38. Wang J, Liu D, Ip WH, Zhang W, Deters R. Integration of system-dynamics, aspect-programming, and object-orientation in System information modeling. *IEEE Trans Ind Inform* (2014) 10(2):847–53. doi:10.1109/tii.2014.2300703

39. Vernotte JF, Panciatici P, Meyer B, Antoine JP, Deuse J, Stubbe M. High fidelity simulation of power system dynamics. *IEEE Computer Appl Power* (1995) 8(1):37–41. doi:10.1109/67.363304

40. Durgaprasad G, Thakur SS. Robust dynamic state estimation of power systems based on M-Estimation and realistic modeling of system dynamics. *IEEE Trans Power Syst* (1998) 13(4):1331–6. doi:10.1109/59.736273

41. Eremia M, Shahidehpour M. Power System modeling and control. In: *Electrical power system dynamics: modeling, stability, and control*. IEEE (2013).

42. Mora M, Cervantes-Pérez F, Gelman-Muravchik O, Forgionne GA. Modeling the strategic process of decision-making support systems implementations: a System dynamics approach review. *IEEE Trans Syst Man, Cybernetics, C (Applications Reviews)* (2012) 42(6):899–912. doi:10.1109/tsmcc.2011.2171482

43. Gu J, Qin Y, Bao Hl., Li XK. Dynamic comprehensive evaluation of distribution network planning based on entropy weight and system dynamics. *Power Syst Prot Control* (2013).

44. Wu Y, Lou S, Lu S. A model for power system interconnection planning under low-carbon economy with CO_2 emission constraints. *IEEE Trans Sustainable Energy* (2011) 2(3):205–14. doi:10.1109/TSTE.2011.2118241

45. Khansari N, Vesaghi A, Mansouri M, Mostashari A. The multiagent analysis of social progress in energy behavior: the system dynamics methodology. *IEEE Syst J* (2017) 11(4):2062–71. doi:10.1109/jsyst.2015.2457771

46. Oluic M, Berggren B, Echavarren FM, Ghandhari M, Rouco L. On the Nature of voltage impasse regions in power System dynamics studies. *IEEE Trans Power Syst* (2018) 33(3):2660–70. doi:10.1109/tpwrs.2017.2763123

47. Fan AJ, Pan YY. The development potential and relative benefit exploration of the global energy interconnection. In: *Fu Jian forum, humanities and social edition* 7 (2016).

48. Forrester JW. Urban dynamics. *IMR; Ind Management Rev (Pre-1986)* (1970) 11(3):67.

49. Forrester JW. System dynamics, systems thinking, and soft OR. *Syst Dyn Rev* (1994) 10(2-3):245–56. doi:10.1002/sdr.4260100211

50. Hou J. The sustainable development of Harbor based on System dynamics. *Syst Eng Theor Pract* (2010) 30(1). doi:10.1016/S1876-3804(11)60009-8

51. Zhang CX, Liu ZX. *The analysis of cultural and tourist industry in Henan Province, China based on system dynamics*. China Academic Journal Electronic Press (2014).

52. Zuo QT, Chen L. Social economic ecological environment coupling System dynamics model. *Shanghai Environ Sci* (2001) 20(12).

53. Yao YP, Yang YX, Sun HY, Yu XF. Research on environmental benefit of recycling mechanical and electrical products based on the normalization method. In: *Mechanical design and manufacturing* (2011).

54. Tan JJ, Wang PP, Li Y. Application of System dynamics in comprehensive benefit evaluation of demand response. *Automation Power Syst* (2014). doi:10.7500/AEPS20130716017

55. He KL. *Study on technical economic analysis and comprehensive benefit evaluation of intelligent power distribution*. Beijing: North China Electric Power University (2016).

56. IRENA. *Renewable energy outlook for ASEAN*. IRENA & ACE (2016).

57. Forman EH, Saul IG. The analytical hierarchy Process—An exposition. *Operations Res* (2001) 49(4):469–87. doi:10.1287/opre.49.4.469.11231

58. Gao CY, Shao H, Bi HP, Zhan ZP, Song XS. System dynamics model of risk management of power construction project. *Oper. Res. Eng. Sci. Theory Appl.* (2017).

59. Yu SQ, Hou CH, Sun YT. Power demand forecast in Shandong Province, China based on system dynamics. *J. Shandong Univ. (Eng. Sci.)* (2015).

frontiers | Frontiers in Sustainable Food Systems

TYPE Original Research
PUBLISHED 18 September 2025
DOI 10.3389/fsufs.2025.1608981

Check for updates

OPEN ACCESS

EDITED BY
Ze Wang,
Capital Normal University, China

REVIEWED BY
Nanfei Jia,
Beijing Technology and Business University, China
Junjie Li,
Institute of Agricultural Resources and Regional Planning (CAAS), China

*CORRESPONDENCE
Catherine Brinkley
✉ ckbrinkley@ucdavis.edu

RECEIVED 09 April 2025
ACCEPTED 26 August 2025
PUBLISHED 18 September 2025

CITATION
Brinkley C, Ulimwengu J and Raj S (2025) Real and simulated shocks to the food system: a nested food-economic system model for assessing resilience using a networked input–output approach.
Front. Sustain. Food Syst. 9:1608981.
doi: 10.3389/fsufs.2025.1608981

COPYRIGHT
© 2025 Brinkley, Ulimwengu and Raj. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Real and simulated shocks to the food system: a nested food-economic system model for assessing resilience using a networked input–output approach

Catherine Brinkley[1]*, John Ulimwengu[2] and Subhashni Raj[3]

[1]Department of Human Ecology, University of California, Davis, Davis, CA, United States, [2]International Food Policy Research Institute, Washington, DC, United States, [3]Department of Urban and Regional Planning, University of Hawaii at Manoa, Honolulu, HI, United States

Introduction: This study provides a novel, system-wide framework that integrates food and economic systems using networked input–output.
Methods: We demonstrate this framework in analyzing food system resilience to real and simulated sector-level shocks, yielding insights into how structural vulnerabilities within the economy can compromise the food system and vice versa. The nested and networked approach of considering the food system as embedded within the overall economic system, enables predictions within and across both systems as they relate to one another. We focus on recovery and resiliency after a shock by evaluating the U.S. food and economic system in 2007 and 2012, a period covering a major financial downturn.
Results and Discussion: We find that between 2007 and 2012, the majority (82.2%) of food systems sectors increased in centrality compared to other sectors within the economy, showing how the food system became more integral in the overall economic system during and after the financial crisis. Further, we demonstrate that simulated removal of sectors with the highest centrality in the economic system leads to a decrease in overall network density, clustering, degree, and weighted degree—highlighting vulnerabilities within each economic sector with the potential to propagate system-wide. We draw timely attention to the critical role of scientific research as an important sector within the economy that requires relatively little funding input but amplifies economic outputs across the food system and other economic sectors. We conclude by warning about the bias produced in analyzing food system dynamics outside of overall economic systems and urge future research to consider a nested food-economic model.

KEYWORDS

food system resilience, networked input–output, nested food-economic resilience model, economic shock, complex systems

Introduction

Why focus on food system resilience?

As election debates continuously highlight, the price of groceries remains a salient motivator at the polls. Urban riots reveal how political combustible any food insecure

populations can be, and political volatility in one city can lead to further destabilization and food insecurity well beyond. One of the key analytical tools for safeguarding food and economic security is to predict areas of vulnerability and their networked, interdependence with the greater universe of allied economic sectors. Predictions of various failure scenarios provide insights into overall resilience and recovery from shocks within particular sectors.

Scholars and policymakers are increasingly recognizing the need to model how shocks propagate across economic and food systems. For its part, the food system encompasses the processes that require inputs to generate outputs involved in feeding a population, including growing, harvesting, processing, packaging, transporting, marketing, consuming, and disposing of food (Hawkes, 2009; Sobal et al., 1998; Ericksen, 2008; Tu et al., 2019). Each aspect of the food system is subject to climate, social, and economic disruption through pricing, policies, or localized disaster. For example, localized failure can result from crop disease (as in the potato famine of 1845) or closure of central processing plants (as in the closure of slaughter houses during COVID-19). Changes induced by various economic, political, climate, markets, and infrastructural factors can also disrupt food supply at multiple levels, from growing food to distributing supplies (Davis et al., 2021; Watson et al., 2017). Even when components of the food system are resilient in the face of different shocks and crises, expected outcomes such as food and nutrition security are not always guaranteed for all (Piters et al., 2021)—a finding we explain is due to the food system being nested within the overall economic system where some communities cannot afford to access the food that is produced and delivered.

Given the increasing occurrence of political, socio-economic, and environmental instability with climate change and extreme weather events (Cottrell et al., 2019), it is critical to understand how food systems react to shocks and stressors, as a food-insecure population can rapidly cause disruption in many other sectors. We define food system resilience as the capacity to withstand shocks or rapidly recover in the face of stressors.

This research advances the field by demonstrating that food system resilience requires conceptualizing the food system as an integral part of the broader economic, political, and environmental systems. The intensification of modern scientific and engineering methods in farming, processing, and distribution technologies further reinforces the connection between the food system and the rest of the economic system (Matson et al., 1997). Yet, a review by Tendall et al. (2015) shows that most studies on food system resilience lack a systems approach, focusing instead on selected segments such as agricultural production or specific stages in the food value chain (Van Apeldoorn et al., 2011; Soane et al., 2012) or particular components of resilience like adaptability and transformability (Walker et al., 2009), cascading effects in regime shifts (Kinzig et al., 2006), or specific outcomes and contexts, e.g., food security in emergency situations (Pingali et al., 2005). Given that resilience varies across locations, time, and scales, vulnerability to shocks should be correspondingly analyzed over time and across food system components (Tendall et al., 2015; Konar et al., 2018).

Measuring the resilience of food systems through simulated shocks is a critical approach to understanding the ability to withstand and recover from various disruptions. By subjecting food systems to simulated shocks, researchers can assess vulnerabilities and identify strategies to enhance resilience. Several studies have employed this methodology, modeling the food system as a network (Karakoc and Konar, 2021; Lin et al., 2019; Gephart and Pace, 2015; Brinkley et al., 2025). Building on such network studies, Burkholz and Schweitzer (2019) explore the impacts of cascading shocks on global crop trade networks from 176 countries across 21 years. Their findings highlight the role of higher order trade interactions, and they note that "networks reveal hidden dependencies between countries" that are likely the result of economic and trade agreements beyond crops (Burkholz and Schweitzer, 2019, p. 1). Similarly, in modeling commodity price shocks related to the Ukraine-Russia war, Laber et al. (2023) demonstrate how a localized shock to agricultural production in one country can create food shortages well beyond immediate trade partners. While such simulated shocks help assess losses from weather events like drought to identify critical vulnerabilities, authors note that networked food supply models often treat each commodity in isolation of the broader food system, neglecting how products may be further processed or used. Along these lines, studies do not always nest the food system within the broader economic system and therefore miss the buffering and propagating impacts of economic sectors that feed into the food system.

Similarly, recent studies leverage input–output (IO) frameworks to quantify food-system resilience under shocks. For example, Revoredo-Giha and Dogbe (2023) applied a dynamic inoperability IO model (DIIM) to Scotland's economy, simulating a ~ 60% demand collapse in hospitality and tracking recovery paths. They computed sectoral "resilience coefficients" and inoperability trajectories, finding that preserved meats and agriculture recovered fastest (high coefficients) while soft drinks and other food processors recovered slowest (coefficients ≈0.1). This dynamic IO approach explicitly models how an initial shock cascades through intermediate demand, incorporates adaptive strategies (e.g., telework or inventory buffers as higher resilience parameters), and yields recovery times and cumulative losses for each sector. Huang et al. (2025) synthesize 26 empirical studies to assess how global food supply systems (FSS) respond to sudden shocks like COVID-19 and natural disasters, emphasizing that systems with strong operational, relational, and structural attributes are best positioned for recovery, adaptation, and transformation. However, they find that while many FSS display short-term adaptability, few convert these coping mechanisms into long-term structural reforms, especially in resource-constrained settings. Rocha Aponte (2024) used Leontief–Ghosh IO analysis on Norway's seafood sector, revealing strong backward/forward linkages and low import dependence. The fisheries/aquaculture sector remained stably connected to food, beverage and tobacco industries even amid COVID, suggesting adaptive supply-chain policies that maintained resilience. Together, these modeling studies—from Scotland to Norway (and by extension U.S. analyses using similar DIIM methods)—show how systemic shocks propagate through food networks. They identify vulnerable sectors and quantify recovery dynamics, informing policies (e.g., diversification or local sourcing) to bolster cross-sector resilience.

In building from these studies, this research combines methods in input–output modeling with a network approach to understand value creation and economic functionality of the food system within the broader economic system. We utilize the concept of the input-supply network, where an economy is described through a set of specialized production units or nodes, each relying on the flow of

inputs from their suppliers to produce their own output, which in turn is routed towards other downstream units, forming a production network (Carvalho and Voigtländer, 2014; Lin et al., 2014). Production units are measured in dollar value. While a production network approach does not include household food consumption, Taschereau-Dumouchel (2019) points out that understanding the structure of the production network is key to determining whether and how microeconomic shocks affecting only a particular node propagate throughout the economy and shape network outcomes. Such shock propagation will ultimately affect the main outcome of the food system, which is food and nutrition security, by impacting the availability and pricing of food. Further, our dataset covers the global food crisis of 2008, when a convergence of different market shocks and disruptions in food production caused dramatic increases in global food prices and food shortages (Headey, 2011).

Input–Output (I-O) captures the interdependence of different sectors in an economic system. This interdependence is critical for understanding how shocks propagate through the system (Miller and Blair, 2009). Several studies have applied I-O models to measure economic resilience. For example, Hallegatte (2008) used an I-O model to simulate the economic impacts of natural disasters, finding that the resilience of an economy largely depends on its ability to substitute inputs, which can be measured using I-O tables. Koks and Thissen (2016) extended this approach by considering the resilience of global supply chains, showing how I-O tables can capture the global propagation of local shocks. Infrastructure systems are also increasingly being studied using I-O models. Haimes (2009) argued that I-O models can measure the resilience of infrastructure systems by capturing their interdependence. This approach was later applied by Ouyang and Dueñas-Osorio (2012), who used an I-O model to measure the resilience of interconnected infrastructure systems after a natural disaster. I-O models have also been used to measure the resilience of ecological systems. Finnoff and Tschirhart (2008) used an I-O model to capture the interaction between economic and ecological systems and measure their resilience. Avelino and Hewling (2019) used I-O tables to examine the resilience and adaptability of regional economies, highlighting the role of economic concentration and activity outflow in shaping regional resilience. Inoue and Todo (2019) used I-O tables to examine how shocks propagate at the firm level through supply-chain networks, providing valuable insights into the resilience of individual firms and their supply chains, and by extension, the resilience of the systems in which they operate. Contreras and Fagiolo (2014) found that (i) the more a sector is globally central in the country network, the larger its impact; and (ii) the largest European countries, such as those constituting the core of the European Union's economy, are more vulnerable to economic shocks. The model developed by Carvalho (2014) explicitly highlights the role of input-supply linkages, where an idiosyncratic shock affecting a single sector is transmitted to its downstream neighbors in the network and, via the latter, propagates further downstream to other production nodes indirectly connected with the original sector.

Despite these promising applications, several challenges remain in using I-O tables to measure system resilience. For one, I-O models are linear, which may not adequately capture the non-linear behavior of real-world systems (Dietzenbacher and Los, 1997). Furthermore, I-O tables are typically static, making it difficult to capture dynamic changes in system resilience (Acemoglu et al., 2012). Addressing these challenges will require the development of more advanced I-O models and the integration of I-O models with other modeling approaches.

To address these shortcomings, we combine the standard input–output (I-O) framework developed by Leontief (1949) with a network approach. In the context of complex networks, by combining food and non-food sectors, the use of I-O tables avoids truncation bias as explicit food system nodes may have multiple linkages with nodes or sectors outside the mainstream food system. For example, chemistry (for fertilizers and pesticides) and machinery (for agricultural equipment) are critical to farming activities. By using a networked I-O approach, we are also able to quantify the bias in more truncated systems measures of resilience that focus only on the food system and not the greater economic system. Studies on other sectors of the economy note that node or sector removal simulations based on highest degree (Albert et al., 2000; Callaway et al., 2000; Cohen et al., 2001) or highest betweenness centrality (Magoni, 2003) have the ability to propagate.

We use a network centrality approach, which allows for multiple features of the system to be observed by implementing node deletion scenarios. Deng et al. (2007) argue that node deletion does not necessarily destroy the connectedness of a growing network if the increasing rate of edges is not excessively small. Early studies on network resilience focusing on node deletion assume that once the system has fallen apart, some key functions or features may be lost (Albert et al., 2000); for example, it is expected that once a critical number of server nodes are down, the internet will collapse. This study explores how the food system operates without the overall economic system and vice versa. We also highlight vulnerable sectors that can drastically change system functionality.

While node deletion simplifies the complex nature of economic systems, it is employed here as a stress-testing tool to identify structurally fragile sectors within the network. Crucially, removing a node inherently eliminates all of its associated edges—effectively severing that sector's connections to the rest of the network. This means that node deletion can also be interpreted as a form of targeted edge deletion, representing the abrupt disruption of critical intersectoral linkages such as supply chain dependencies or technological collaborations. From this perspective, the approach offers a useful proxy for simulating realistic economic shocks without the need to model every potential mode of partial failure explicitly. This interpretation addresses concerns about the realism of node deletion by situating it within a broader framework of network disruption analysis.

Methods

Any system can be represented as a network of graphs where a given graph G is a pair of sets (V,E) of nodes (or agents) V and edges (or relationship) E with their cardinality denoted by $|V|$ and $|E|$, respectively. Weighted and oriented networks can be described using an adjacency matrix $=(w_{ij})$, $\forall i,j=1,\ldots,n$, where w_{ij}represents the level of interaction between nodes i and j. In our study, each edge in the network represents an actual financial transaction between two sectors, as reported in the U.S. Input–Output (I-O) tables. Specifically, the weight of a directed edge from sector i to sector j corresponds to the dollar value of intermediate goods or services that sector i provides

to sector j.[1] This approach allows us to simulate the structural importance of sectors based on their actual economic footprint, rather than relative proportions, thereby offering a more realistic view of systemic vulnerability and resilience.

It follows that a network can also be defined by $N=(V,W)$. A generic k-path with starting node $i_0 \in V$ is given by $p_{i_0}^{(k)} = \{i_0 \to i_1 \to \cdots \to i_{k-1} \to i_k\}$, where $i_0, i_1, \cdots, i_{k-1}, i_k$ are distinct nodes in V and arcs $i_{h-1} \to i_h \in E, \forall h = 1, \ldots, k$.

All the k-paths of network N starting from node i_0 can be assembled as shown in Equation (1):

$$P^{(k)} = \bigcup_{i_0 \in V} p_{i_0}^{(k)} \tag{1}$$

Following Carvalho (2014), Acemoglu et al. (2016) and Acemoglu and Guerrieri (2008), in Equation (2), we assume that each node of the network is approximated by the following constant return to scale technology:

$$X_i = b_i A_i^{\beta} l_i^{1-\beta} \left(\prod_{j=1}^{n} X_{ij}^{w_{ij}} \right)^{\beta} \tag{2}$$

where A_i is the corresponding productivity shock, the amount of production input used by node i is given by l_i, while $1-\beta$ is the share of input in production; b_i is a given constant; X_{ij} denotes the amount of good from node j supplied to node i while the exponent $w_{ij} (\geq 0)$, element of adjacency matrix W, represents the share of good from node j in the total intermediate input used by node i.

In this context, the adjacency matrix W is defined by three main elements (Carvalho, 2014): (i) a collection of n nodes, each node corresponding to one of the sectors in the economy; (ii) a collection of directed edges, where an edge between any two nodes denotes an input-supplying relationship between two nodes; and (iii) a collection of weights, each of which is associated with a particular directed edge and given by the exponent w_{ij} in the output function.

We assume a representative consumer with Cobb–Douglas preferences, as shown in Equation (3):

$$u(c_1, \ldots, c_n) = \tilde{b} \prod_{i=1}^{n} c_i^{\frac{1}{n}} \tag{3}$$

where c_i is the amount of good i consumed and $\tilde{b}$ is some positive constant.

The first-order conditions of the firm's problem at node i are given in Equations (4) and (5), representing input allocation and labor choice, respectively:

$$X_{ij} = \beta w_{ij} p_i X_i / p_j \tag{4}$$

$$l_i = (1-\beta) p_i X_i / \omega \tag{5}$$

where ω and p_i represent, respectively, the wage market and the price of good i. Since $X_i = c_i + \sum_{j=1}^{n} X_{ij}$, it follows that the optimal production for node i is given by Equation (6)

$$X_i = b_i A_i^{\beta} \phi_i (1-\beta)^{1-\beta} \prod_{j=1}^{n} \left(\beta w_{ij} X_i / \phi_j \right)^{\beta w_{ij}} \tag{6}$$

for some constant ϕ_j.

Given a constant b_i, Taking the log-linear transformation of the production function yields Equations (7) and (8), simplifying downstream aggregation:

$$x_i = \beta \sum_{j=1}^{n} w_{ij} x_j + \beta \epsilon_i, \text{or} \tag{7}$$

$$x = \left[1 - \beta W \right]^{-1} \epsilon \tag{8}$$

where $\epsilon_i = \log(A_i)$.

As pointed out by Acemoglu et al. (2017), because of constant return to scale assumption, nodes make zero profits in equilibrium; hence, value added is simply equal to the market wage ω. Assuming $\frac{n}{\tilde{b}} (p_1 \ldots p_n)^{\frac{1}{n}} = 1$ with $p_i = \phi_i \omega / X_i$, the log real value of the economy is derived in Equation (9), capturing system-wide output adjusted for productivity.

$$\log \omega = \frac{1}{n} \sum_{i=1}^{n} \log(X_i) - \frac{1}{n} \sum_{i=1}^{n} \log(\phi_i) + \log\left(\frac{\tilde{b}}{n} \right) \tag{9}$$

For an appropriate choice of $\tilde{b}$, as shown in Equation (10), aggregate output is computed as the mean of node-level logs:

$$y = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{10}$$

As shown in Equation (11), without loss of generality, Carvalho (2014) shows that aggregate output, y, is simply a weighted sum of the (logarithm of) nodes-level shocks, ϵ_i:

$$y = \sum_{i=1}^{n} v_i \epsilon_i \tag{11}$$

where the weights, v_i, are determined by the network matrix W, $v = \frac{1-\beta}{n} (I - \beta W')^{-1} 1$.

Equation 11 states that aggregate output might fluctuate over time due to shocks from individual nodes; the magnitude of these aggregate fluctuations can be traced back to the network structure through adjacency matrix W using network centrality measures such as the Katz–Bonacich centrality measure.

The Katz–Bonacich centrality vector is expressed in Equation (12), which links adjacency matrix W to centrality θ. The Katz–Bonacich eigenvector centrality measure is derived by assigning to each node j, a centrality weight $\theta_j > 0$, which is defined by some baseline centrality level, equal across all nodes, plus a term which is proportional to the

1 This differs from traditional input–output technical coefficients, which normalize these flows by the total inputs or outputs of a sector. In contrast, we preserve the raw monetary transaction values to capture the true magnitude and strength of interdependencies in the economic network. These transaction-based weights are directly embedded in the adjacency matrix W, which is used throughout the analysis.

weighted sum of the centrality weights of its downstream nodes: $\theta_j = \lambda \sum_i w_{ij}\theta_i + \text{ł}$ for some parameter $\lambda > 0$. In matrix form, this implies that the vector of centralities is given by:

$$\theta = \eta\left(I - \lambda W'\right)^{-1} 1 \tag{12}$$

where θ is the vector of Katz–Bonacich centralities given an input–output network W where $\eta = \frac{1-\beta}{n}$; $\lambda = \beta$, the share of intermediate inputs in Equation 2.

It follows that Equation 11 can be written as

$$y = \sum_{i=1}^{n} \theta_i \epsilon_i \tag{13}$$

Hence, Equation 13 establishes the link between aggregate network outcome and its centralities. Finally, Acemoglu et al. (2016) show that in any given interaction network, node or agent *i*'s Bonacich centrality can be written as

$$\theta_i = 1 + \beta \sum_{j=1}^{n} \theta_j w_{ij} \tag{14}$$

Equation 14 implies that node *i* has a higher centrality (and hence a more pronounced impact on the rest of the nodes) if it interacts strongly (if β is close to 1) with nodes that are themselves central. In other words, "a shock to an agent with a higher Bonacich centrality impacts the states of other relatively central agents, which in turn propagates the shock further to other agents, and so on, eventually leading to a larger aggregate impact" (Acemoglu et al., 2016, p. 586). Therefore, network centrality measures can be used to assess overall system resilience.

Finally, network volatility can also be measured by

$$\sigma_y = \sigma_{\text{ò}} \sqrt{\sum_{i=1}^{n} \theta_i^2} \tag{15}$$

Equation 15 can be decomposed into economy-wide sub-systems which include the food system.

While complete sectoral collapse is rare, node deletion serves as a stylized tool for stress-testing the network, identifying structurally critical sectors, and mapping the propagation of failures. The method is conceptually aligned with existing literature in complex networks and systemic risk modeling (e.g., Albert et al., 2000; Acemoglu et al., 2016), where node or edge deletion simulates extreme disruptions like strikes, natural disasters, or policy shutdowns. This modeling experiment provides insights into fragility and interdependence that real-world variability might obscure.

Data and metrics

The empirical implementation of a simple analytical framework laid out in the previous section is conducted by using 2007 and 2012 Input–Output Tables (IOTs) of the United States of America (USA) from the Bureau of Economic Analysis (BEA) as networks that describe, for a given period, transactional flows within an economy. The 2007 and 2012 time points were selected to capture pre- and post-2008 financial crisis dynamics. This interval allows examination of how economic shocks affect intersectoral dependencies and which sectors gained or lost influence as a result. This implies that that economic shocks—such as the 2008 financial crisis—impact not only the volume of sectoral production but also the structural position of sectors within the broader economic network. In other words, beyond measurable changes in output, shocks can reconfigure intersectoral dependencies by altering how sectors are interconnected and how influential they are within the system.

The US has I-O tables over 1997–2021 period but with a maximum of 71 industries compared to 405 industries for 2007 and 2012. In addition to their disaggregated structures, the choice of 2007 and 2012 I-O tables allows for an assessment of system resilience after the 2008 food and financial shocks. The estimates are available at four levels of detail (see Howells et al., 2018): sector (21 industry groups), summary (71 industry groups), underlying summary (138 industry groups), and detail (405 industry groups). Sectors are presented according to the 2012 North American Industry Classification System (NAICS) code structure. In this paper, after cleaning data to remove negative entries and isolated nodes, the resulting network comprises 363 nodes and 43,599 edges for 2007, and 362 nodes and 42,906 edges for 2012.[2] In a fully connected network where every node/sector has a connection to all other nodes, the overall system would have 131,406 edges for the 2007 sector (N*(N−1)). Thus, the US economic system has many, differently weighted, interdependencies, but not every sector is connected to every other sector, opening interesting opportunities to model resiliency. To comply with network analysis requirement, the diagonal of adjacency matrix W was set to zero for each network.

Following the North American Industry Classification System (NAICS) (Horowitz and Planting, 2006), the selection of food sector includes processing (for example, flour milling and malt

2 In constructing the input–output networks for 2007 and 2012 from the U.S. Bureau of Economic Analysis (BEA) detailed I-O tables (comprising 405 sectors), we undertook the following data cleaning steps:

Removal of negative entries: Negative transaction values (arising from adjustments, returns, or imputed transactions in I-O tables) were excluded, as they can distort the representation of positive economic flows in a network model.

Exclusion of isolated nodes: Nodes (sectors) that had no incoming or outgoing financial transactions (i.e., rows and columns with all zeros) were removed, as they represent non-participating sectors within the active economic network in the specified year.

Food industry nodes: Using NAICS codes, we identified 43 nodes as food-system related in both 2007 and 2012 networks. This classification includes agricultural production, food processing, distribution, retail, and related services (e.g., restaurants).

These steps ensure an analytically tractable and representative economic network where each node has economic activity. Removing inactive or anomalous nodes improves network robustness and the interpretability of results, especially under simulations like node deletion.

TABLE 1 List of scenarios and deleted nodes.

Node deletion scenario	Nodes deleted
Removal of top 10 sectors with highest centrality measure	1. Scientific research and development services 2. Other amusement and recreation industries 3. Nonresidential maintenance and repair 4. Architectural, engineering, and related services 5. Civic, social, professional, and similar organizations 6. Limited-service restaurants 7. Accommodation 8. Full-service restaurants 9. Management of companies and enterprises 10. Other motor vehicle parts manufacturing
Removal of bottom 10 sectors with lowest centrality	1. Carbon and graphite product manufacturing 2. Independent artists, writers, and performers 3. Electric lamp bulb and part manufacturing 4. Household refrigerator and home freezer manufacturing 5. Manufacturing and reproducing magnetic and optical media 6. Automatic environmental control manufacturing 7. Support activities for agriculture and forestry 8. Photographic and photocopying equipment manufacturing 9. Audio and video equipment manufacturing 10. Electronic computer manufacturing
Removal of top 10 food sectors with highest centrality measure	1. Limited-service restaurants 2. Full-service restaurants 3. All other food and drinking places 4. Food and beverage stores 5. Grocery and related product wholesalers 6. Bread and bakery product manufacturing 7. Fruit and vegetable canning, pickling, and drying 8. Soft drink and ice manufacturing 9. Cookie, cracker, pasta, and tortilla manufacturing 10. Fluid milk and butter manufacturing
Removal of bottom 10 food sectors with lowest centrality	1. Wet corn milling 2. Flour milling and malt manufacturing 3. Flavoring syrup and concentrate manufacturing 4. Fruit and tree nut farming 5. Soybean and other oilseed processing 6. Support activities for agriculture and forestry 7. Greenhouse, nursery, and floriculture production Support activities for agriculture and forestry 8. Fishing, hunting and trapping 9. Coffee and tea manufacturing 10. Beef cattle ranching and farming, including feedlots and dual-purpose ranching and farming

manufacturing) and non-processing (for example, grain farming) food activities, and related supporting activities/services (for example, full-service restaurants, farm machinery and equipment manufacturing, and support activities for agriculture and forestry). To assess the resilience of US food system to shocks, we run a series of simulations by deleting nodes with the highest degree centrality. Furthermore, the volatility of the system is assessed by using a measure developed by Acemoglu et al. (2016, equation 15 in Supplementary material). We look also at the impact of shocks in the form of nodes deletion on the network outcome (here, production). Finally, we assess the bias of analyzing food system resilience based on truncated food system as in most previous studies. We define a truncated food system as a network of food sectors disconnected from non-food sectors in the overall economic system. In a way, food system truncation corresponds to an extreme case of the deletion of non-food nodes. The complete list of deleted nodes for each scenario is presented in Table 1.

A brief description of network metrics in the analysis is provided in in Table 2. These metrics include betweenness centrality, clustering coefficient, density and total degree centrality.

Results

Network centrality

We find that over the study period, the U.S. food system consolidated nearly twice as much as the overall economic system. In 2007, the US economic network had a density (the number of existing

TABLE 2 Selected network metrics.

Network metric	Definition and application to US trade input–output matrix
Betweenness centrality	Betweenness centrality captures a node's role in allowing information to pass from one part of the network to the other. It captures how much a given node is in-between others. This metric is measured with the number of shortest paths (between any couple of nodes in the graphs) that passes through a given node.
Clustering coefficient	In many studies, a network's clustering coefficient was identified as a good indicator for network robustness. It captures how nodes are embedded in their respective neighborhoods. First proposed by Watts and Strogatz (1998), it is defined as the average of the local clustering coefficient of its nodes. A node's clustering coefficient measures how close its neighborhood is to a complete network in terms of the relative density of links in its neighborhood. Our findings confirm that node deletion may decrease the network's robustness.
Density	By design, the top 10 sectors (overall and food) are sets of central and densely connected nodes, while the bottom 10 are noncentral and sparsely connected nodes, namely, the periphery. In a way, the top 10 sectors belong to the "rich-club," a set in which the highest degree nodes show a high density of connections. It then follows that the removal of central nodes will likely decrease the density of a network and ultimately affect its resilience. Similarly, because of their sparseness, removal of noncentral nodes will likely increase the density of a network by reducing the size of its periphery; therefore, reinforcing its resilience.
Eigenvector centrality	Eigenvector centrality is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. In this paper, we use Katz-Bonacich centrality for eigenvector centrality
Total degree centrality	The degree of a node is the number of incoming and outgoing relations (edges) the node has; it is weighted by the value of each edge. Contend that resilience centrality is mainly determined by the degree and weighted nearest-neighbor degree of the node, in which weighted nearest-neighbor degree plays a prominent role.

connections relative to the potential connections (PC[3]) of 0.332 compared to 0.328 in 2012; in other words, between the two periods, the US economy lost 1.6% of its edges and 1.2% of its density. The food subnetwork (Figure 1) also saw a reduction in density from 0.357 in 2007 to 0.346 in 2012, losing 3% of its edges and density over the two time points. Such consolidation can be protective against shock or give added exposure to shock depending on network realignments and dependencies, as we describe next.

We also observed significant shifts in the centrality rankings of sectors, providing insights into their relative systemic importance. Specifically, 58% of sectors increased in centrality, indicating that they became more embedded and influential within the economic network (Figures 2–5). In contrast, 37% of sectors experienced a decline in their centrality rankings, suggesting reduced integration or weakening connections to the rest of the economy. These changes highlight that sectors do not respond uniformly to shocks; instead, some become more dominant while others diminish in systemic relevance. For example, the "Internet publishing and broadcasting" sector rose by 92 ranks, while "Newspaper publishing" dropped 61 places, showing how the shock reshaped economic hierarchies. The maximum increase in ranking was 92 spots ("Internet publishing and broadcasting and Web search portals") and the largest decrease was 96 rank spots ("Broadcast and wireless communications equipment"). Thus, "intersectoral dependencies" are not static—they evolve with macroeconomic events, and our centrality analysis captures that dynamic.

Notably, the top 25 sectors in the economy did not experience a large degree of change in either rank or eigenvector centrality (Figure 4). The apparent stability of these top sectors in face of the 2008–2009 food and financial crisis may be a sign of the US economy capacity to absorb such shock. Only the "Scientific research and development services" sector remained unchanged and held onto its ranking as the most central sector in the economic system. Interestingly, despite the overall decline in eigen centrality, 62% of nodes in the economic system saw an increase in its production between 2007 and 2012.

Like the overall economic system, the food system experienced significant movement in ranking and centrality between 2007 and 2012 (Figure 6). Though unlike the economic system (Figure 4), the change in ranking among the food sectors was minor Figure 5). While 42% (18 nodes) of the food sector nodes experienced an increase in ranking (Figure 5), the range of rank increase was 3, followed by no change in ranking (32% of nodes), and only a quarter of nodes (11) experienced a decline in ranking. Again, the decline was relatively small compared to rank changes in the economic system (range of 10, Figures 2–4). Changes in rank are commensurate with changes in eigenvector centrality. Of the food sectors, "Breakfast cereal manufacturing" achieved the largest gain in eigen centrality, increasing 14.2% between 2007 and 2012 (Figure 5). In contrast, "food and beverage stores" sector saw the largest slump, dropping 44.3% in eigen centrality between 2007 and 2012 (Figure 5). Followed closely by the "grocery and related product wholesalers" sector experiencing a 41.9% reduction in their eigen centrality over the same period (Figure 5). With the financial crisis and less disposable income available, food purchasing and thereby sectors involved in food retail became less central to the economy (Figure 6 c-d). Despite these notable exceptions, 60% of the 43 food sector nodes increased in (eigen) centrality between 2007 and 2012 (Figures 5 and 6c-d) The production value[4] of 88% of the food sectors experienced an increase during the same period (Figures 5 and 6c-d). Between 2007 and 2012, 24 food

3 PC = n*(n–1)/2, where n is the number of nodes in the network.

4 Production value, is the total value of goods and services measured in USD.

10.3389/fsufs.2025.1608981

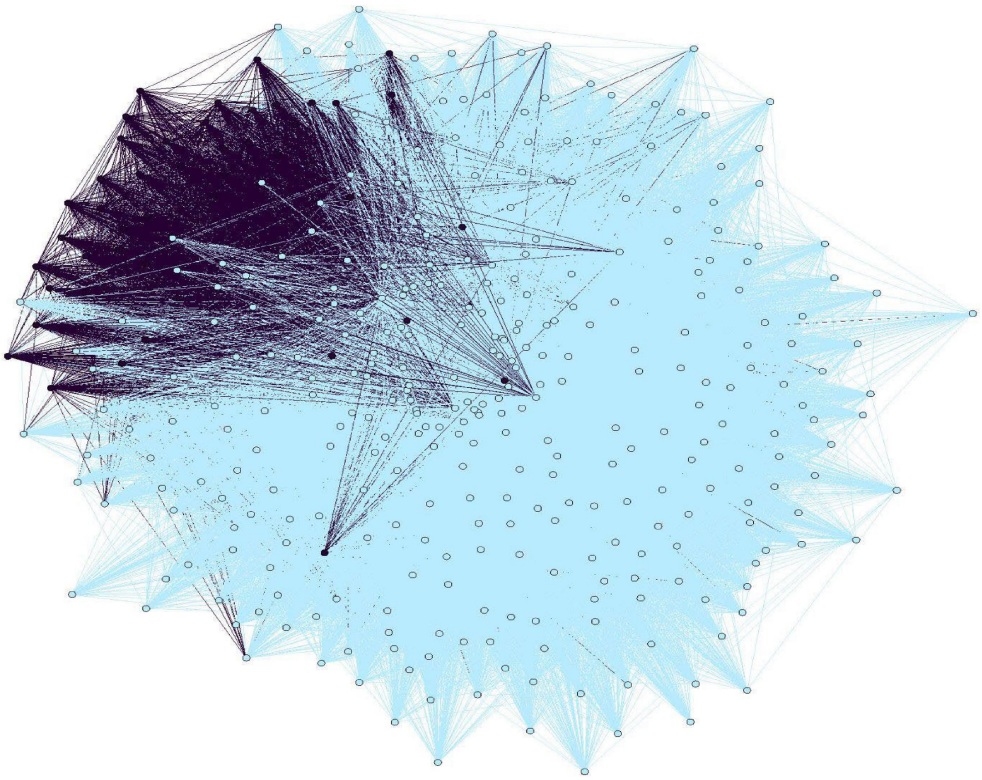


FIGURE 1
Network of the overall US economic system nodes and edges (light blue) with the food system nodes and edges (dark blue). Note that some more central nodes in the food system are also more central to the overall economic system (2007 I-O data).

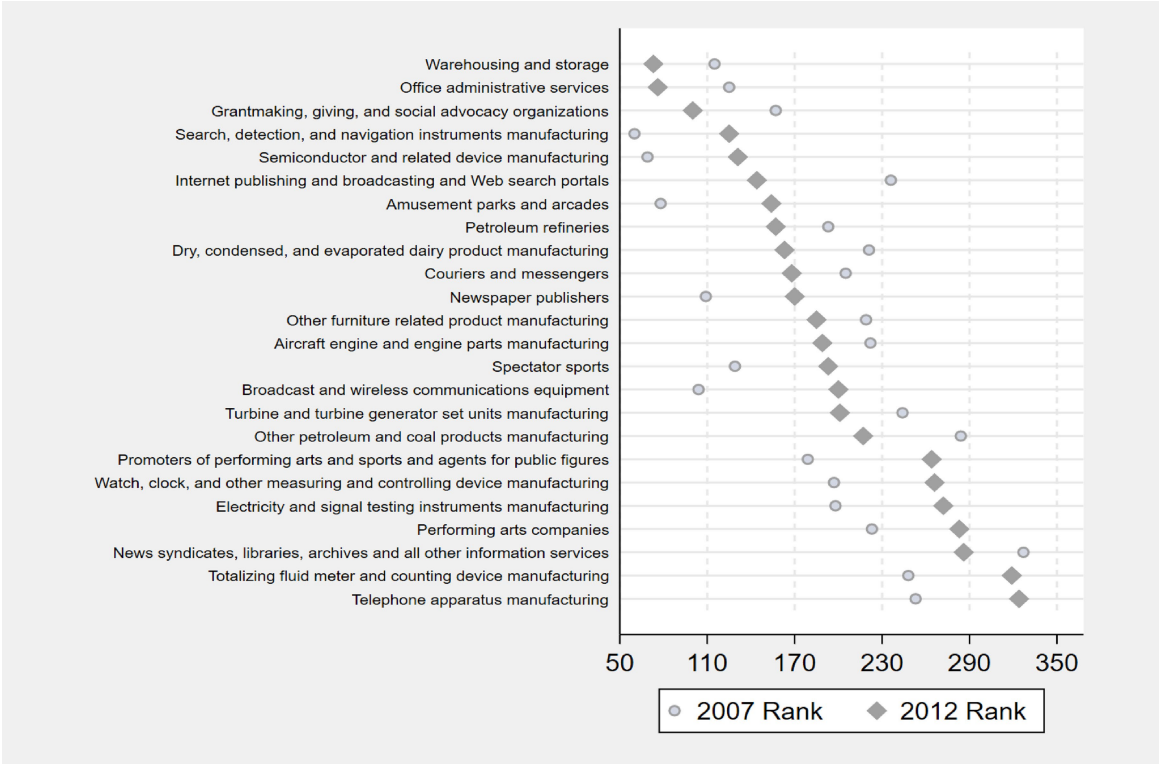


FIGURE 2
Sectors with the greatest rank change from 2007 to 2012.

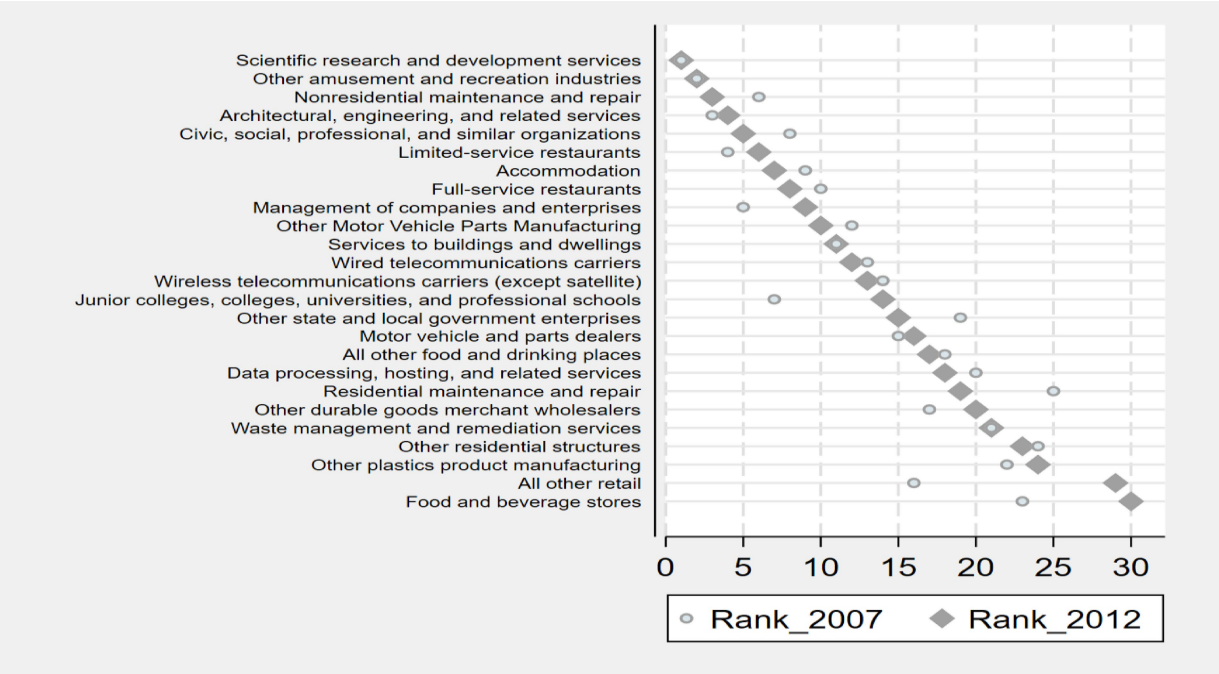


FIGURE 3
Change in ranking for the top 25 economic sectors between 2007 and 2012.

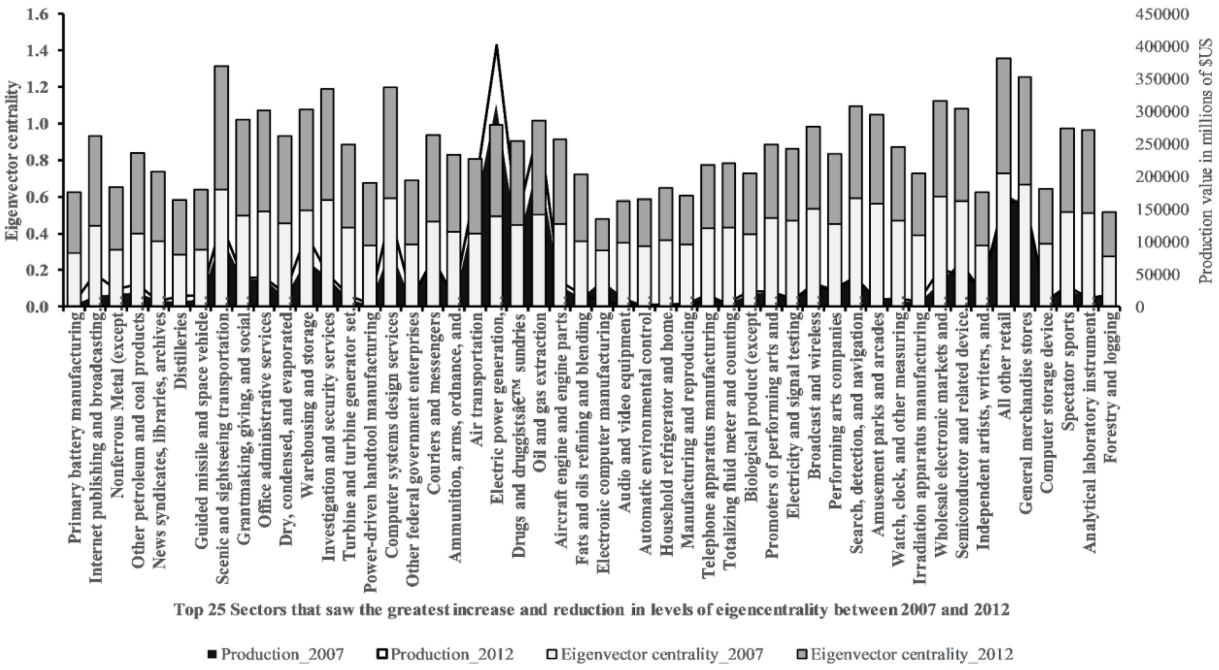


FIGURE 4
Economic sectors with the greatest increases (25 in total) and decreases (25 in total) in eigen centrality from 2007 to 2012.

sectors experienced an increase in both production value and in centrality (Figures 5 and 6c-d).

Overall, as shown in Figures 6a,b, we find a positive correlation between sector centrality and production as derived in Equation 13. This is an important feature as a negative shock that affects a central sector is likely to reduce its production and propagate the shockwave throughout the network by creating an aggregate disruption. As pointed out by Carvalho (2014), in a

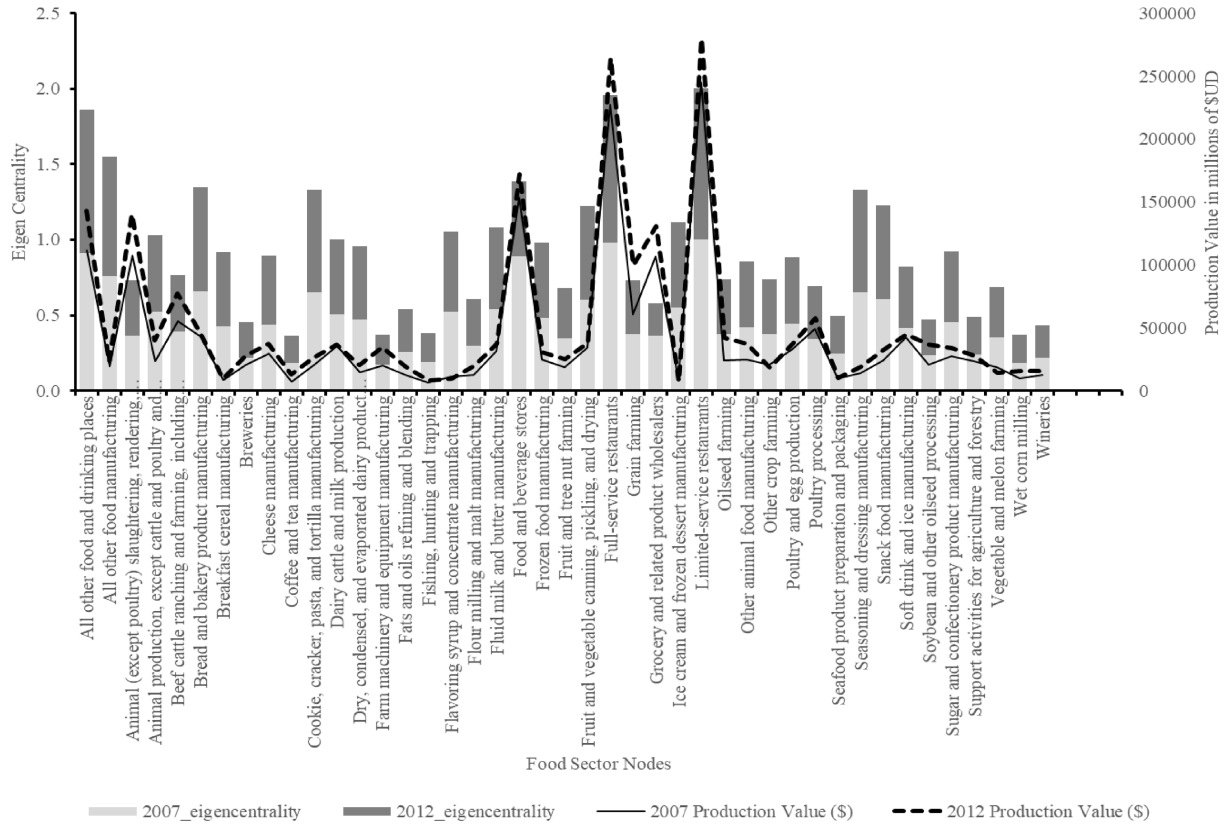


FIGURE 5
Change in eigen-centrality and production value in the food sector between 2007 and 2012.

production network that stresses the role of input-supply linkages, "an idiosyncratic shock affecting a single sector will be transmitted to its downstream neighbors in the network and, via the latter, propagate further downstream to other production nodes which are only indirectly connected with the original sector" (p. 38). Hence, the resilience of a given food system is more likely to rest on the sectors that exhibit strong correlation between production and centrality.

Figure 7 presents the level of volatility[5] with respect to centrality between the two periods. The results suggest that the volatility is much higher for the overall economic system compared to the food system. Given that the overall economic system has more nodes compared to the food system, the law of large numbers, where aggregate output becomes increasingly concentrated around a constant value, does not hold here. As pointed out by Gabaix (2011), when some sectors play a disproportionately important role (due to their much larger size relative to others), the standard central limit theorem no longer holds. As a result, aggregate fluctuations may result from sector-level (idiosyncratic) shocks (Acemoglu et al., 2017). Over time, though not significant, volatility was on average slightly lower in 2012 compared to 2007 which might be related to the recovery from the 2008–2009 crisis as mentioned earlier. It is probable that the higher volatility in both the networks in 2007 is indicative of the global financial crisis that preceded the 2008 crisis.

Sector failure scenarios

Highly centralized networks, like the U.S. economic system, rely heavily on the most central nodes for overall connectivity and functionality such that impacts on these nodes will have cascading effects across the network. We show that deletion of the bottom 10 economic and food nodes tends to boost density, clustering, degree, and weighted degree of the overall economic network (see Figures 8a–c). While deletion of the most central 10 economic nodes negatively impacts the same network characteristics.

We also find that betweenness centrality is affected but has a reverse relationship with the node deletion scenarios described above indicating that influences on the many smaller sectors that act as bridges to other sectors are more likely result in cascading effects. Both deletion scenarios demonstrate different types of fragility in the overall food and total economic system. In general, we expect that the impact of a shock to a system defined by a network of interconnected nodes would not be linear as the direction of the impact clearly depends on the place of each node in the network. However, the larger point remains that a shock in the form of node deletion is likely to disrupt the system resilience and ultimately its outcome.

As presented in Table 3, overall, the results show that dropping a node (sector) is likely to decrease the system's overall outcome (as

5 Metric defined in Equation 14 assuming = 1 (see Supplementary material).

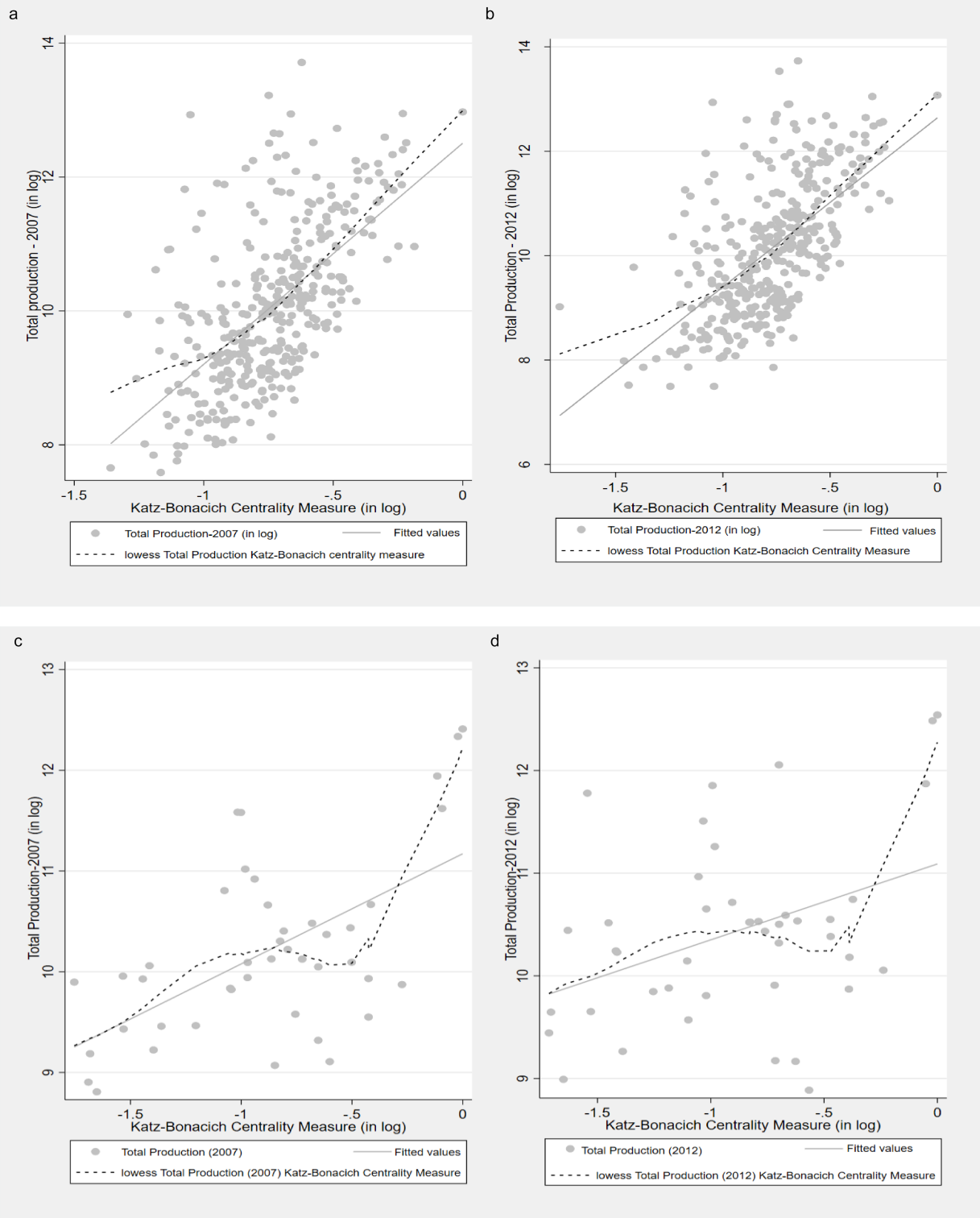


FIGURE 6
(a) Production and centrality for the overall economy in 2007. **(b)** Production and centrality for the overall economy in 2012. **(c)** Production and centrality for the food system in 2007. **(d)** Production and centrality for the food system in 2012.

measured in Equation 13). Compared to the baseline, deleting the bottom 10 economic nodes decreased the production outcome of the overall system by 0.9 percent. When considering only the food sectors, removing the top food nodes produced the largest drop in production outcome: 3.6% demonstrating the overall economic system's reliance on the food system. In comparison, deletion of the bottom 10 food nodes led to a 0.75 percent decrease in the production outcome.

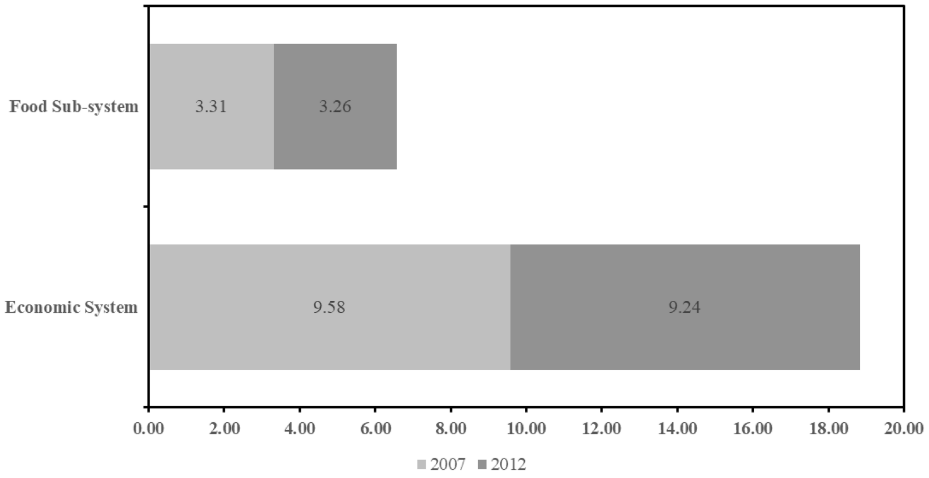


FIGURE 7
Volatility of food and overall economic systems between 2007 and 2012.

Paradoxically, we also find that deleting the top 10 economic nodes *increased* production outcome of the overall network relative to baseline by 27% (Table 3). In disaggregating these sectors, we show that only the "Scientific research" sector is related to such increases in production, with a 24% increase. This sector is also the only top 10 sector whose *outdegree* value is higher than its *indegree* value, likely because relatively few "research and development" dollars can result in major innovation and economic returns. In this sense, "Scientific research" is an amplifier of economic returns for the entire system. In explanation, deleting central nodes in a network can disrupt the flow of funding, isolating sections of the network, and making them appear as though the remaining nodes are receiving more input and producing more output than before.

Impact of evaluating the food system in isolation

Next, we compare the impact of analyzing resilience based on a truncated food system to further demonstrate the value of considering food system dynamics more wholistically within the full economic system. We present the truncated food system in Table 4, showing the impact of removing non-food sectors from the economy. Using the US 2012 I-O table, truncation leads to a reduction of the weighted values of incoming and outgoing connections by 53.0 and 21.6%, respectively. This is a dramatic change in food system network functionality would render inaccurate any finding drawn from evaluating the food system in isolation of the rest of the economic system. The most affected sectors include "All other food and drinking places," "Breweries," "Farm machinery and equipment manufacturing," "Food and beverage stores" and "Full-service restaurants".

As expected, truncation also affects the clustering of the food system which is reduced by 22.5% on average. Except for "Wineries" whose clustering coefficient increases by 15.1%, all the other sectors experience reduction in their clustering coefficient. The top five sectors most affected include: "Fruit and tree nut farming" (−33.8%), "Fruit and vegetable canning, pickling, and drying" (−34.1%), "Animal (except poultry) slaughtering, rendering, and processing" (−34.2%), "Soybean and other oilseed processing" (−34.3%), and "Wet corn milling" (−38.0%). Such changes in clustering and centrality likely bias predictions of overall food system function when considering the food system as separate from the overall economic system.

The most striking effect of truncation is on centrality values. As discussed earlier, centrality is used to identify sectors which are critical for both the network performance and resilience. Table 4 shows how the ranking of sectors with respect to their central role in the network shifted as a result of truncation. Out of 43 food sectors, only 5 kept their central ranking ("Limited-service restaurants," "Full-service restaurants," "All other food and drinking places," "Fluid milk and butter manufacturing," and "Fishing, hunting and trapping"), 18 experienced a decrease in their rankings (for example, "Grocery and related product wholesalers" moved from 5th to 39th position), and 20 climbed in ranking (for example, "Flavoring syrup and concentrate manufacturing" moved from 37th to 12th position). In short, by considering *only* the food sector, centrality measures become unreliable in terms of capturing the key features and functionality of the food network—including measuring its resilience. These findings add caution to any interpreting the results of any study that uses an incomplete economic system to evaluate performance.

Discussion and conclusion

The primary objective of this research is to assess the resilience of the overall food system to cope with shocks in any of its sectors by considering the food system as nested within the broader network of the overall economic system. Our findings offer caution when interpreting findings from studies that analyze food system dynamics in isolation from the economic system. Specifically, we demonstrated that truncating the food system reduced the weighted values of incoming and outgoing connections by 53.0 and 21.6%, respectively. Moreover, the centrality rankings of food sectors shifted due to truncation, demonstrating that even centrality measures become unreliable when not considering the broader economic context. Such substantial changes render inaccurate any findings if the food system functionality is analyzed in isolation. Future research should take heed

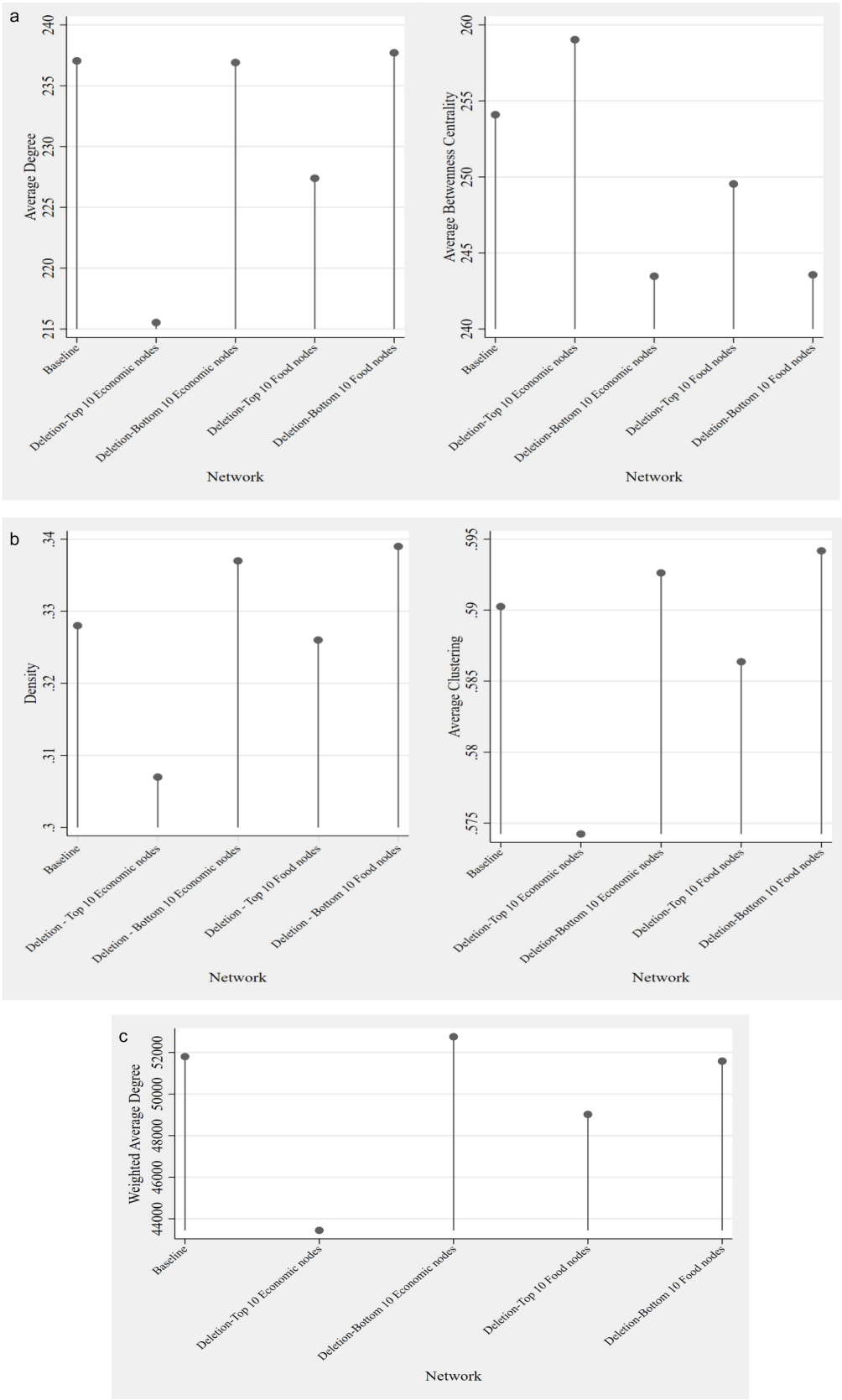


FIGURE 8
(a) Impact of node deletion on the average degree (left) and average betweenness centrality (right) of the overall economic network. **(b)** Impact of node deletion on the density(left) and average clustering (right) of the overall economic network. **(c)** Impact of node deletion on the average weighted degree of the overall economic network.

TABLE 3 Impact of nodes deletion on production of overall economic system.

Network	Scenario (node deletion)	Production	Percent change
1	Baseline	171.118	
2	Top 10 economic sectors	217.537	27.13
3	Bottom 10 economic sectors	169.547	−0.92
4	Top 10 food sectors	164.956	−3.60
5	Bottom 10 food sectors	169.824	−0.76
6	Scientific Research	211.896	23.83
7	Accommodations	169.582	−0.90
8	Architectural Services	169.458	−0.97
9	Civic	169.746	−0.80
10	Full service restaurant	169.459	−0.97
11	Management Service	169.599	−0.89
12	Limited service restaurant	169.471	−0.96
13	Non-residential	169.384	−1.01
14	Other amusement	169.936	−0.69
15	Other motor	169.888	−0.72

and construct models that connect the food system to the broader economic system.

We also draw the conclusion that the resilience of the food system enhances the resilience of the overall economic system based on several strands of empirical evidence. Most notably, between 2007 and 2012—a period that includes the global financial crisis—82.2% of food-related sectors increased in centrality. This indicates that food sectors became more influential and better connected within the broader economic network, suggesting their growing importance in sustaining intersectoral flows during a time of systemic stress. We also show that the volatility of the overall economic system was higher than that of the food system during the 2007–2008 financial crisis, indicating a buffering effect within the nested food system sectors. The lower volatility in the food system compared to the broader economy suggests that the U.S. food system is a stabilizing force in the overall economy. Rather than amplifying volatility, as some central sectors might, the food system appears to have acted as a buffer, helping to preserve the continuity of production across the economic network. Further supporting this conclusion is the simulation of node deletion scenarios. While the most central food sectors are unlikely to ever be fully removed from the economy in a real-world scenario, the simulated removal of the most central food-related sectors—those involved in food retail and processing, for example—generated a 3.6% decline in overall system output. This impact was significantly larger than the decline resulting from the removal of the least central food sectors (0.75%) or the bottom-ranked economic sectors more broadly (0.9%). The fact that the removal of key food sectors resulted in a more substantial systemic loss to the overall economic system underscores the food system's structural importance and the extent to which other sectors depend on it for their own functioning. This finding is crucial for modeling food system resilience in other countries where economic contexts differ and reliance on food systems for core economic functions may be even greater.

Our findings imply that maintaining food system functionality during crises is not only essential for food security but also contributes to overall macroeconomic stability. Taken together, these findings offer robust evidence that the food system contributes positively to systemic resilience. Its rising centrality during a crisis, the disproportionate impact of its disruption, and its relative stability all point to its essential role not just in food security, but in safeguarding macroeconomic functionality during and after major shocks. Such findings support the prioritization of food system investments in economic recovery planning.

More broadly, our research also found a positive correlation between sector centrality and production, implying that shocks reducing the production of central sectors are likely to propagate and cause system-wide disruption. For example, we show that between 2007 and 2012, the network centrality measures of many sectors changed due to the financial shock; 37% (135) of economic sectors experienced a drop in centrality rankings, while 58% (211) saw an increase. For instance, the centrality increase of food sectors like "breakfast cereal manufacturing" post-2008 mirrors increased consumer demand for lower-cost staples during economic downturns. Conversely, the decline in centrality of retail nodes aligns with income constraints affecting food purchases. Interestingly, the top 25 sectors in the economy remained relatively stable in their centrality rankings, suggesting either the resilience of these core sectors or that they are foundational to the U.S. economy and were thus buffered from the shock. As various sectors within the food system and broader economy continued to consolidate (e.g., Hendrickson et al., 2001; Hendrickson et al., 2020; Howard, 2014; Howard, 2015), networked I-O modeling techniques like those we demonstrate in this research can help show where such centralization will be stabilizing or destabilizing under a variety of shock scenarios. Future studies can also assess how emergency policies that support essential sectors help mitigate shock propagation or further spur centralization and vulnerability.

To this end, our results confirm that the removal of nodes with high centrality measures leads to a decrease in network metrics such as density, clustering, degree, and weighted degree, emphasizing the importance of central nodes in maintaining overall system resilience. These findings underscore the often-overlooked role of natural

TABLE 4 Comparing 2012 food network measures between full and truncated food systems.

Sectors[a]	Weighted indegree			Weighted outdegree			Clustering			Eigen centrality		
	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)
All other food and drinking places	78,756	8,487	−89.2	39,351	1946	−95.1	0.4372	0.3417	−21.9	0.6983	0.9500	36.0
All other food manufacturing	16,668	10,496	−37.0	4,442	3,541	−20.3	0.6448	0.4318	−33.0	0.4767	0.7880	65.3
Animal (except poultry) slaughtering, rendering, and processing	100,515	81,769	−18.6	21,221	13,848	−34.7	0.6465	0.4253	−34.2	0.4382	0.3708	−15.4
Animal production, except cattle and poultry and eggs	9,175	4,205	−54.2	28,042	26,330	−6.1	0.5916	0.4160	−29.7	0.3808	0.5116	34.4
Beef cattle ranching and farming, including feedlots and dual-purpose ranching and farming	31,776	14,335	−54.9	56,325	56,063	−0.5	0.6643	0.5147	−22.5	0.3086	0.3746	21.4
Bread and bakery product manufacturing	28,459	16,625	−41.6	3,200	2,193	−31.5	0.7094	0.5433	−23.4	0.5240	0.6887	31.4
Breakfast cereal manufacturing	7,708	5,250	−31.9	20	10	−50.0	0.7678	0.5917	−22.9	0.4103	0.4892	19.2
Breweries	18,219	4,938	−72.9	7,645	4,947	−35.3	0.7285	0.6091	−16.4	0.4661	0.2421	−48.0
Cheese manufacturing	24,656	19,779	−19.8	8,900	8,533	−4.1	0.7244	0.6275	−13.4	0.4751	0.4543	−4.4
Coffee and tea manufacturing	10,383	6,697	−35.5	2,441	1,583	−35.1	0.8158	0.6964	−14.6	0.3132	0.1802	−42.5

(Continued)

10.3389/fsufs.2025.1608981

TABLE 4 (Continued)

Sectors[a]	Weighted indegree			Weighted outdegree			Clustering			Eigen centrality		
	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)
Cookie, cracker, pasta, and tortilla manufacturing	19,079	12,205	−36.0	1713	1,581	−7.7	0.7064	0.4915	−30.4	0.5075	0.6785	33.7
Dairy cattle and milk production	29,584	15,382	−48.0	36,157	36,149	0.0	0.6022	0.4723	−21.6	0.4256	0.4974	16.9
Dry, condensed, and evaporated dairy product manufacturing	17,403	11,753	−32.5	4,568	3,641	−20.3	0.6961	0.5541	−20.4	0.4753	0.4876	2.6
Farm machinery and equipment manufacturing	24,107	49	−99.8	4,190	3,301	−21.2	0.6500	0.5275	−18.9	0.5631	0.1959	−65.2
Fats and oils refining and blending	11,231	9,349	−16.8	9,489	6,925	−27.0	0.6932	0.4815	−30.5	0.3631	0.2855	−21.4
Fishing, hunting and trapping	2,413	91	−96.2	10,252	9,581	−6.5	0.6400	0.5192	−18.9	0.3180	0.1921	−39.6
Flavoring syrup and concentrate manufacturing	4,619	2,532	−45.2	11,748	11,253	−4.2	0.7158	0.5216	−27.1	0.3566	0.5348	50.0
Flour milling and malt manufacturing	16,548	12,307	−25.6	14,088	13,672	−3.0	0.7347	0.5163	−29.7	0.3598	0.3057	−15.0
Fluid milk and butter manufacturing	27,372	19,746	−27.9	5,115	4,650	−9.1	0.7155	0.5736	−19.8	0.5066	0.5393	6.5

(Continued)

10.3389/fsufs.2025.1608981

TABLE 4 (Continued)

Sectors[a]	Weighted indegree			Weighted outdegree			Clustering			Eigen centrality		
	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)
Food and beverage stores	69,481	2050	−97.0	2,257	1,186	−47.5	0.5535	0.4400	−20.5	0.6307	0.4970	−21.2
Frozen food manufacturing	19,979	14,706	−26.4	2,632	2,282	−13.3	0.6964	0.4801	−31.1	0.4763	0.4969	4.3
Fruit and tree nut farming	8,313	2,959	−64.4	16,382	16,354	−0.2	0.6416	0.4249	−33.8	0.3546	0.3316	−6.5
Fruit and vegetable canning, pickling, and drying	22,971	13,363	−41.8	4,814	4,211	−12.5	0.6832	0.4501	−34.1	0.5210	0.6234	19.6
Full-service restaurants	114,701	23,067	−79.9	50,341	1909	−96.2	0.3263	0.3188	−2.3	0.7426	0.9789	31.8
Grain farming	63,300	11,871	−81.2	76,002	55,777	−26.6	0.6206	0.4292	−30.8	0.4174	0.3559	−14.7
Grocery and related product wholesalers	59,026	488	−99.2	48,234	39,558	−18.0	0.3655	0.3362	−8.0	0.5785	0.2135	−63.1
Ice cream and frozen dessert manufacturing	4,874	2,965	−39.2	1715	1,492	−13.0	0.7189	0.5789	−19.5	0.4200	0.5677	35.2
Limited-service restaurants	145,793	30,743	−78.9	18,577	799	−95.7	0.3337	0.3182	−4.6	0.7651	1.0000	30.7
Oilseed farming	17,410	3,547	−79.6	22,593	22,020	−2.5	0.6167	0.5026	−18.5	0.4019	0.3608	−10.2
Other animal food manufacturing	33,703	24,771	−26.5	32,711	32,060	−2.0	0.7112	0.4921	−30.8	0.4476	0.4363	−2.5
Other crop farming	9,082	2069	−77.2	9,316	7,328	−21.3	0.5741	0.4167	−27.4	0.4031	0.3606	−10.6
Poultry and egg production	30,022	17,744	−40.9	27,727	27,662	−0.2	0.6342	0.4383	−30.9	0.3727	0.4377	17.5

(Continued)

10.3389/fsufs.2025.1608981

TABLE 4 (Continued)

Sectors[a]	Weighted indegree			Weighted outdegree			Clustering			Eigen centrality		
	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)	Full	Truncated	Change (%)
Poultry processing	38,751	29,259	−24.5	7,740	7,319	−5.4	0.7540	0.6667	−11.6	0.4502	0.3486	−22.6
Seafood product preparation and packaging	7,951	5,801	−27.0	3,892	2,941	−24.4	0.7747	0.6439	−16.9	0.3762	0.2495	−33.7
Seasoning and dressing manufacturing	14,027	7,246	−48.3	5,876	4,903	−16.6	0.6872	0.4723	−31.3	0.4538	0.6768	49.2
Snack food manufacturing	22,718	11,260	−50.4	3,534	3,281	−7.2	0.6812	0.5127	−24.7	0.4747	0.6243	31.5
Soft drink and ice manufacturing	34,696	9,722	−72.0	4,080	3,150	−22.8	0.7315	0.7182	−1.8	0.5208	0.4044	−22.4
Soybean and other oilseed processing	28,817	23,650	−17.9	27,726	25,175	−9.2	0.6448	0.4234	−34.3	0.3406	0.2343	−31.2
Sugar and confectionery product manufacturing	18,418	7,082	−61.5	8,300	7,938	−4.4	0.6594	0.4484	−32.0	0.4835	0.4673	−3.3
Support activities for agriculture and forestry	6,437	724	−88.8	25,625	21,097	−17.7	0.6706	0.4762	−29.0	0.3219	0.2433	−24.4
Vegetable and melon farming	6,522	1,476	−77.4	2,693	2,581	−4.2	0.6875	0.5261	−23.5	0.3718	0.3333	−10.4
Wet corn milling	10,571	7,680	−27.3	13,459	9,561	−29.0	0.7040	0.4368	−38.0	0.3605	0.1818	−49.6
Wineries	10,029	3,040	−69.7	4,003	2,947	−26.4	0.8211	0.9500	15.7	0.4144	0.2167	−47.7

[a]For comparison purpose, we kept only food sectors.

10.3389/fsufs.2025.1608981

systems, upon which the food system and economic system draws. Like many studies, this study did not incorporate the many unpriced natural sectors that underpin the overall food and economic system. Should our model be expanded to integrate natural systems into I-O models, the overall assessment could enhance the accuracy of resilience results and help to more accurately model environmental degradation externalities across economic sectors.

On the one hand, high-degree nodes might increase the resilience of a network because these nodes can disperse information or influence to many other nodes in the network. For example, a study by Albert et al. (2000) found that the Internet's resilience to random failures was due to its scale-free topology, characterized by a few hubs with high outdegree values that can "pick up the slack" when other clusters of nodes fail. On the other hand, these nodes could also decrease network resilience if targeted themselves. If a node with a high outdegree fails or is removed, this could disrupt a significant proportion of the network, as many nodes might depend on the high outdegree node. This idea is supported by a study by Callaway et al. (2000), who found that removing high-degree nodes could cause a rapid disintegration of the network. Nodes with higher outdegree than indegree also play a crucial role in cascading failures. Further, a study by Motter and Lai (2002) demonstrated that the failure of a single node can lead to a cascade of failures in other nodes, particularly in scale-free networks with high-degree nodes. Finally, high out-degree nodes also have a substantial impact on the diffusion process in the network. For instance, Kempe et al. (2003) demonstrated how more central nodes significantly affect the speed and reach of diffusion in social networks. In sum, nodes with a higher outdegree value than an indegree value can both increase and decrease network resilience, depending on the context.

This study uncovered the critical role of "Scientific research" in both the overall economic system and the food system. This sector buffers from shocks in other sectors, but when deleted itself can result in a cascading failure. The finding that a simulated deletion of "Scientific research" results in an overall increase in economic productivity of 24% can be explained with an analogy. If one were to think of funding towards "Scientific research" as the budget for a maintenance department that takes care of a rental property, it is easy to understand how firing the maintenance staff would reduce costs and produce a higher return on investments. Yet, such a decision would ultimately devalue the assets and cut into long term costs. In this sense, as political budget-making leans towards austerity measures, cuts to higher education and similar research sectors may achieve greater economic outcomes, but with depreciation of both the value of the underlying system and its longer-term potential returns.

This research has important implications for economic policy and resilience planning. We demonstrate how policymakers could design strategies that enhance the resilience of both food and non-food sectors, recognizing their interdependence. Going forward, our nested food-economic systems model can aid predictive modeling to test the impacts of localized shocks. For instance, during the COVID-19 pandemic, labor shortages in U.S. meat packing houses likely impacted several other economic sectors. Using our model, such vulnerabilities could be anticipated and managed more effectively during crises. Investments in statistical systems are necessary to build such comprehensive databases that capture the full spectrum of participants and activities within the food system. Identifying vulnerable sectors, populations, and supply chain locations can help target interventions more effectively.

While the primary empirical focus of this study centers on the period surrounding the 2008 financial crisis, the nested food-economic system model is designed for broader application across a wide range of systemic disruptions. Recent events—such as the Ukraine war's impact on global grain markets (FAO, 2024; Glauber and Laborde Debucquet, 2023), the 2023–2024 El Niño-induced agricultural volatility, and the Red Sea shipping crisis disrupting critical trade flows (Glauber and Mamun, 2024)—illustrate the continuing relevance of a network-based resilience framework. These shocks reinforce the importance of analyzing food system resilience within the broader economic structure, as disruptions in non-food sectors (e.g., shipping, energy, fertilizer) can propagate into food systems with severe consequences. Thus, the insights and simulations in this study are not only historically grounded but also have direct applicability for current and future crisis scenarios. Additionally, countries with significant informal economies likely rely on truncated economic models, leading to biased results similar to those demonstrated in our model of the truncated food system.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

CB: Conceptualization, Writing – review & editing, Funding acquisition, Writing – original draft, Supervision. JU: Conceptualization, Validation, Writing – original draft, Methodology. SR: Validation, Writing – review & editing, Methodology, Data curation, Visualization, Investigation, Conceptualization.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. CB and SR were supported by the National Science Foundation CAREER Award (#1750621), "Mapping pathways to food security and sustainable development."

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2017)."The network origins of aggregate fluctuations," in *The Oxford handbook of the economics of networks*. eds. Y. Bramoullé, A. Galeotti and B. Rogers (Oxford, UK: Oxford University Press). doi: 10.1093/oxfordhb/9780199948277.013.24

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica* 80, 1977–2016. doi: 10.3982/ECTA9623

Acemoglu, D., and Guerrieri, V. (2008). Capital deepening and nonbalanced economic growth. *J. Polit. Econ.* 116, 467–498. doi: 10.1086/589523

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. Networks, shocks, and systemic risk. *The Oxford Handbook of the Economics of Networks* 568–608 (2016).

Albert, R., Jeong, H., and Barabási, A. L. (2000). Error and attack tolerance of complex networks. *Nature* 406, 378–382. doi: 10.1038/35019019

Avelino, A. F. T., and Hewings, G. J. D. (2019). The Challenge of Estimating the Impact of Disasters: Many Approaches, Many Limitations and a Compromise. *Advances in Spatial and Economic Modeling of Disaster Impacts*. Advances in Spatial Science. eds. Y. Okuyama and A. Rose (Cham: Springer). doi: 10.1007/978-3-030-16237-5_7

Brinkley, C., Ulimwengu, J., and Raj, S. Real and simulated shocks to the food system: a nested food-economic system model for assessing resilience using a networked Input-Output approach. *Frontiers in Sustainable Food Systems*, 9:1608981. doi: 10.3389/fsufs.2025.1608981

Burkholz, R., and Schweitzer, F. (2019). International crop trade networks: the impact of shocks and cascades. *Environmental research letters*, 14, 114013. doi: 10.1088/1748-9326/ab4864

Callaway, D. S., Newman, M. E., Strogatz, S. H., and Watts, D. J. (2000). Network robustness and fragility: percolation on random graphs. *Phys. Rev. Lett.* 85, 5468–5471. doi: 10.1103/PhysRevLett.85.5468

Carvalho, V. M. (2014). From micro to macro via production networks. *J. Econ. Perspect.* 28, 23–48. doi: 10.1257/jep.28.4.23

Carvalho, V., and Voigtländer, N. (2014). Input diffusion and the evolution of production networks. doi: 10.3386/w20025

Cohen, R., Erez, K., ben-Avraham, D., and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Phys. Rev. Lett.* 86, 3682–3685. doi: 10.1103/PhysRevLett.86.3682

Contreras, M. G., and Fagiolo, G. (2014). Propagation of economic shocks in input-output networks: a cross-country analysis. *Phys. Rev. E* 90:062812. doi: 10.1103/PhysRevE.90.062812

Cottrell, R. S., Nash, K. L., Halpern, B. S., Remenyi, T. A., Corney, S. P., Fleming, A., et al. (2019). Food production shocks across land and sea. *Nat Sustain* 2, 130–137. doi: 10.1038/s41893-018-0210-1

Davis, K. F., Downs, S., and Gephart, J. A. (2021). Towards food supply chain resilience to environmental shocks. *Nature Food* 2, 54–65. doi: 10.1038/s43016-020-00196-3

Deng, K., Zhao, H., and Li, D. (2007). Effect of node deleting on network structure. *Phys. A Stat. Mech. Appl.* 379, 714–726. doi: 10.1016/j.physa.2007.02.039

Dietzenbacher, E., and Los, B. (1997). Structural decomposition techniques: sense and sensitivity. *Econ. Syst. Res.* 25, 307–323. doi: 10.1080/09535319800000023

Ericksen, P. J. (2008). Conceptualizing food systems for global environmental change research. *Glob. Environ. Chang.* 18, 234–245. doi: 10.1016/j.gloenvcha.2007.09.002

FAO (2024). The global report on food crises 2024. Global Report on Food Crises (GRFC) 2024 | World Food Programme

Finnoff, D., and Tschirhart, J. (2008). Linking dynamic economic and ecological general equilibrium models. *Resour. Energy Econ.* 30, 91–114. doi: 10.1016/j.reseneeco.2007.08.005

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica* 79, 733–772. doi: 10.3982/ECTA8769

Gephart, J. A., and Pace, M. L. (2015). Structure and evolution of the global seafood trade network. *Environ. Res. Lett.* 10:125014. doi: 10.1088/1748-9326/10/12/125014

Glauber, J., and Mamun, A. (2024) Impacts of Red Sea shipping disruptions on global food security. IFPRI blog: issue post. markets, trade, and institutions

Glauber, J. W., and Laborde Debucquet, D. (Eds.) (2023). *The Russia-Ukraine conflict and global food security*. Washington, DC: International Food Policy Research Institute (IFPRI).

Haimes, Y. Y. (2009). On the complex definition of risk: A systems-based approach. *Risk Analysis: An International Journal*, 29, 1647–1654.

Hallegatte, S. (2008). An adaptive regional input-output model and its application to the assessment of the economic cost of Katrina. *Risk Anal.* 28, 779–799. doi: 10.1111/j.1539-6924.2008.01046.x

Hawkes, C. (2009). Identifying innovative interventions to promote healthy eating using consumption-oriented food supply chain analysis. *J. Hunger Environ. Nutr.* 4, 336–356. doi: 10.1080/19320240903321243

Headey, D. (2011). Rethinking the global food crisis: the role of trade shocks. *Food Policy* 36, 136–146. doi: 10.1016/j.foodpol.2010.10.003

Hendrickson, M., Heffernan, W. D., Howard, P. H., and Heffernan, J. B. (2001). Consolidation in food retailing and dairy. *Br. Food J.* 103, 715–728. doi: 10.1108/00070700110696742

Hendrickson, M. K., Howard, P. H., Miller, E. M., and Constance, D. H. (2020). The food system: Concentration and its impacts. A Special Report to the Family Farm Action Alliance, vol. 26, pp. 1–26.

Horowitz, K. J., and Planting, M. A. (2006) Concepts and methods of the us input-output accounts. No. 0066. Bureau of Economic Analysis.

Howard, P. H. (2015). Intellectual property and consolidation in the seed industry. *Crop Sci.* 55, 2489–2495. doi: 10.2135/cropsci2014.09.0669

Howard, P. H. (2014). "Too big to ale? Globalization and consolidation in the beer industry" in *The geography of beer: Regions, environment, and societies*. eds. M. Patterson and N. Hoalst-Pullen (Dordrecht: Springer Netherlands), 155–165. doi: 10.1007/978-94-007-7787-3_14

Howells, T. F. III, Morgan, E. T., and Ross, C. W. (2018). Improved Estimates of the Industry Economic Accounts: Results of the 2018 Comprehensive Update. *Survey of Current Business*, 98.

Huang, I. Y., Forero, O. A., Wagner-Medina, E. V., Florez Diaz, H., Tremma, O., Fargetton, X., et al. (2025). Resilience of food supply systems to sudden shocks: a global review and narrative synthesis. *Glob. Food Secur.* 44. doi: 10.1016/j.gfs.2025.100842

Inoue, H., and Todo, Y. (2019). Firm-level propagation of shocks through supply-chain networks. *Nat. Sustain.* 2, 841–847. doi: 10.1038/s41893-019-0351-x

Karakoc, D. B., and Konar, M. (2021). A complex network framework for the efficiency and resilience trade-off in global food trade. *Environ. Res. Lett.* 16:105003. doi: 10.1088/1748-9326/ac1a9b

Kempe, D., Kleinberg, J., and Tardos, É. (2003). Maximizing the spread of influence through a social network. In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining (137–146).

Kinzig, A. P., Ryan, P., Etienne, M., Allison, H., Elmqvist, T., and Walker, B. H. (2006). Resilience and regime shifts: assessing cascading effects. *Ecol. Soc.* 11:20. doi: 10.5751/ES-01678-110120

Koks, E. E., and Thissen, M. (2016). A multiregional impact assessment model for disaster analysis. *Econ. Syst. Res.* 28, 429–449. doi: 10.1080/09535314.2016.1232701

Konar, M., Lin, X., Ruddell, B., and Sivapalan, M. (2018). Scaling properties of food flow networks. *PLoS One* 13:e0199498. doi: 10.1371/journal.pone.0199498

Laber, M., Klimek, P., and Bruckner, M. (2023). Shock propagation from the Russia–Ukraine conflict on international multilayer food production network determines global food availability. *Nat Food* 4, 508–517. doi: 10.1038/s43016-023-00771-4

Leontief, W. (1949). Structural matrices of national economies. *Econometrica* 17:273. doi: 10.2307/1907314

Lin, X., Dang, Q., and Konar, M. (2014). A network analysis of food flows within the United States of America. *Environ. Sci. Technol.* 48, 5439–5447. doi: 10.1021/es500471d

Lin, X., Ruess, P. J., Marston, L., and Konar, M. (2019). Food flows between counties in the United States. *Environ. Res. Lett.* 14:084011. doi: 10.1088/1748-9326/ab29ae

Magoni, D. (2003). Tearing down the internets. *IEEE J. Sel. Areas Commun.* 21, 949–960. doi: 10.1109/JSAC.2003.814364

Matson, P. A., Parton, W. J., Power, A. G., and Swift, M. J. (1997). Agricultural intensification and ecosystem properties. *Science* 277, 504–509. doi: 10.1126/science.277.5325.504

Miller, R. E., and Blair, P. D. (2009). *Input-output analysis: foundations and extensions*. Cambridge, UK: Cambridge University Press.

Motter, A. E., and Lai, Y. C. (2002). Cascade-based attacks on complex networks. *Phys. Rev. E* 66:065102. doi: 10.1103/PhysRevE.66.065102

Ouyang, M., and Dueñas-Osorio, L. (2012). Time-dependent resilience assessment and improvement of urban infrastructure systems. *Chaos* 22:033122. doi: 10.1063/1.4737204

Pingali, P., Alinovi, L., and Sutton, J. (2005). Food security in complex emergencies: enhancing food system resilience. *Disasters* 29, S5–S24. doi: 10.1111/j.0361-3666.2005.00282.x

Piters, B., Termeer, E., Bakker, D., Fonteijn, H., and Brouwer, H. (2021) Food system resilience towards a joint understanding and implications for policy. Wageningen economic research | policy paper

Revoredo-Giha, C., and Dogbe, W. (2023). A resilience analysis of the contraction of the accommodation and food service sector on the Scottish food industry. *Front. Sustain. Food Syst.* 7. doi: 10.3389/fsufs.2023.1095153

Rocha Aponte, F. (2024). Economic interdependencies and sectoral impacts: an input–output analysis of the Norwegian fisheries and aquaculture sector amidst COVID 19. *Fish. Res.* 281:107247. doi: 10.1016/j.fishres.2024.107247

Soane, I. D., Scolozzi, R., Gretter, A., and Hubacek, K. (2012). Exploring panarchy in alpine grasslands: an application of adaptive cycle concepts to the conservation of a cultural landscape. *Ecol. Soc.* 17. doi: 10.5751/ES-05085-170318

Sobal, J., Kettel Khan, L., and Bisogni, C. (1998). A conceptual model of the food and nutrition system. *Soc. Sci. Med.* 47, 853–863. doi: 10.1016/S0277-9536(98)00104-X

Taschereau-Dumouchel, M. (2019). Cascades and fluctuations in an economy with an endogenous production network. *The Review of Economic Studies*, 2025:rdaf036. doi: 10.1093/restud/rdaf036

Tendall, D. M., Joerin, J., Kopainsky, B., Edwards, P., Shreck, A., Le, Q. B., et al. (2015). Food system resilience: defining the concept. *Glob. Food Secur.* 6, 17–23. doi: 10.1016/j.gfs.2015.08.001

Tu, C., Suweis, S., and D'Odorico, P. (2019). Impact of globalization on the resilience and sustainability of natural resources. *Nat. Sustain.* 2, 283–289. doi: 10.1038/s41893-019-0260-z

Van Apeldoorn, D. F., Kok, K., Sonneveld, M. P. W., and Veldkamp, T. A. (2011). Panarchy rules: rethinking resilience of agroecosystems, evidence from Dutch dairy-farming. *Ecol. Soc.* 16. Avaliable online at: https://www.jstor.org/stable/26268873

Walker, B., Vo, I. D., Beecher, M., and Clancy, C. (2009) A demonstration of the mesh test wireless testbed. 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops, Washington, DC, USA, pp. 1.

Watson, P., Cooke, S., Kay, D., Alward, G., and Morales, A. (2017). A method for evaluating the economic contribution of a local food system. *J. Agric. Resour. Econ.* 42, 180–194. Available online at: https://www.jstor.org/stable/44329749

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442. doi: 10.1038/30918

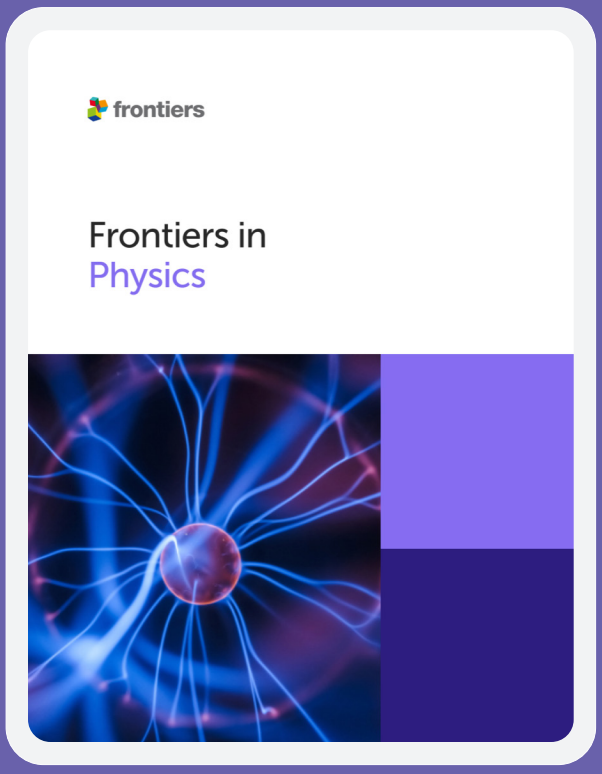
Frontiers in Physics

Investigates complex questions in physics to understand the nature of the physical world

Addresses the biggest questions in physics, from macro to micro, and from theoretical to experimental and applied physics.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact