China Journal of Accounting Research

Volume 18, 3 (2025)

Website:
http://www.elsevier.com/locate/cjar
http://www.cb.cityu.edu.hk/research/cjar/

18
3

China Journal of Accounting Research Vol. 18/3 (2025)

ELSEVIER

Volume 18 • Issue 3 • September 2025

China Journal *of* Accounting Research

HOSTED BY
ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

CYAN MAGENTA YELLOW BLACK

Aims and Scope

The focus of the *China Journal of Accounting Research* is to publish theoretical and empirical research papers that use contemporary research methodologies to investigate issues about accounting, corporate finance, auditing and corporate governance in the Greater China region, countries related to the Belt and Road Initiative, and other emerging and developed markets. The Journal encourages the applications of economic and sociological theories to analyze and explain accounting issues within the legal and institutional framework, and to explore accounting issues under different capital markets accurately and succinctly. The published research articles of the Journal will enable scholars to extract relevant issues about accounting, corporate finance, auditing and corporate governance related to the capital markets and institutional environment.

Production and hosting by Elsevier

Radarweg 29, 1043 NX Amsterdam, The Netherlands

ISSN 1755-3091

Founded by Sun Yat-sen University and City University of Hong Kong

Sponsored by:

Published quarterly in March, June, September, and December

http://www.elsevier.com/journals/china-journal-of-accounting-research/1755-3091/guide-for-authors

CHINA JOURNAL OF ACCOUNTING RESEARCH

Volume 18/3 (2025)

Available online at www.sciencedirect.com

ScienceDirect

HOSTED BY

ELSEVIER

China Journal of Accounting Research 18 (2025) 100431

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Ancient notching or tokening as bookkeeping – Evidence from wood slips in China during 217–210 BCE

Min Chen [a], Li Huang [b,*], Oliver Zhen Li [c], Chunlong Zhang [d]

[a] *Hunan University, China*
[b] *Antai College of Economics and Management, Shanghai Jiao Tong University, China*
[c] *National University of Singapore, Singapore*
[d] *Research Institute of Cultural Relics and Archaeology of Hunan Province, China*

ARTICLE INFO

Article history:
Received 5 April 2025
Accepted 14 May 2025
Available online 2 June 2025

Keywords:
Accounting history
Accounting slips
Notching
Tokening
Ancient accounting

ABSTRACT

We examine Qin Dynasty (221 to 207 BCE) wood bookkeeping slips from 217 to 210 BCE unearthed in Hunan province, China. While purely written slips were unearthed before, these slips are unique in that they contain written records of details of transactions as well as notches on their sides that represent the quantities of money or measures of commodities involved. Scholars have speculated that ancient engraving, notching or tokening before the development of written language could be bookkeeping/accounting. We show a form of bookkeeping combining notching with written records that emerged at a point in Chinese history where a region saw a temporary dip in local literacy. Notching reappeared to compensate for the loss in literacy to prevent fraud and reduce bookkeepers' risk of being accused of fraud. This form of dual-method bookkeeping adds credence to the conjecture that prehistorical, pre-language notching, engraving or tokening was likely bookkeeping/accounting.

1. Conjecture of prehistorical notching or tokening as bookkeeping

Exploring the origin and the evolution of language, researchers have been paying attention to patterns etched on bones and stones in caves as well as unearthed clay tokens and other artifacts. Archeologists believe

* Corresponding author.
E-mail addresses: cmin@hnu.edu.cn (M. Chen), lhuang-acem@sjtu.edu.cn (L. Huang), bizzhenl@nus.edu.sg (O.Z. Li), zclhnca613@sina.com (C. Zhang).

https://doi.org/10.1016/j.cjar.2025.100431
1755-3091/

that they are an early form of recording by human beings dedicated to describing certain objects, food or tools. For example, they show that the Blombos Cave engravings in South Africa dated about 70,000 to 77,000 years ago are intentional images, suggesting that *Homo sapiens* was behaviorally modern around that time, at least in South Africa (Henshilwood et al., 2002; Henshilwood et al., 2018). There are many discoveries similar to the Blombos Cave engravings from other parts of the world dated to different periods in time. For example, the engraved stones dated about 30,000 years ago found in the Shuidonggou Ruin in Northwest China (Peng et al., 2012), the Ishango bone discovered in the Democratic Republic of Congo as well as many more specimens of paleolithic engraved and carved artifacts (Marshack, 1979; 1996; Hovers, 1990; Mackay and Welz, 2008).

Archeologists have argued and attempted to provide evidence that these engravings are unlikely to be randomly caused by nature or human activities. In fact, many have conjectured that they are an early form of recording by human beings deliberately dedicated to describing certain objects or numbers (Peng et al., 2012). Specifically, some archeologists and accounting historians believe that at least some of these etching and drawings represent an early form of bookkeeping/accounting and that bookkeeping or the recording of numbers preceded the development of writing or language. A well-known example is the Middle East "clay token bookkeeping" in ancient Mesopotamia from 8,000 BCE to 3,100 BCE. Clay tokens were baked and sealed inside hollow clay envelops (balls) or impressed on clay balls' exterior or tablets and various shapes and forms represent different objects (Schmandt-Besserat, 1983). This kind of "bookkeeping" (Mattessich, 1987; 1994; Basu and Waymire, 2006) appeared to precede Cuneiform writing. Patterns similar in spirit to these Middle East clay tokens were also found in many other parts of the world. While some researchers believe that ancient bookkeeping preceded abstract counting and written language, whether these tokens or etchings actually represent bookkeeping or accounting/counting is subject to speculations.

2. Notched bookkeeping slips unearthed in Liye Town

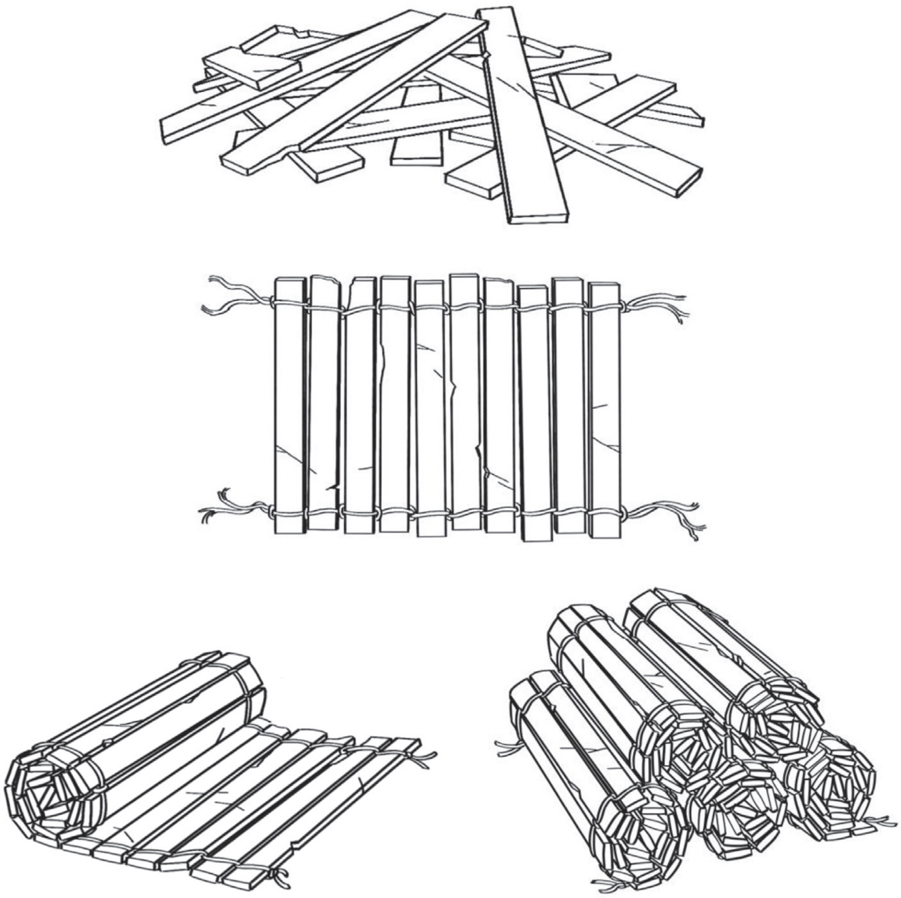
In 2002, a batch of wood slips were unearthed in Liye Town (里耶), Longshan County of Hunan Province in China. Bamboo or wood slips, before the invention and wide-spread use of paper, were a medium for writing and recording in ancient China. Bamboo or wood could be cut into thin slips and these slips could be organized into rolls or scrolls with thin ropes to accommodate a large content (Exhibit 1). During the Warring States Era (475 to 221 BCE), slips were a major medium for writing. By the end of the Jin Dynasty (265 to 420), they yielded to the widespread use of paper.

These unearthed slips were made in 217–210 BCE, during China's Qin (秦) Dynasty (221 to 207 BCE) when China first became a unified state. These slips have ancient Chinese characters written on them, enabling us to understand their content (Exhibit 2). Along their edges, these slips have carved notches of various shapes. There are about 120 of these slips and they are related to bookkeeping as they were dealing with quantities of money or commodities.

As these slips contain both written text as well as carved notches, archeologists (Zhang, et al., 2015) are able to determine that these notches are not random but have numerical and unit of measurement meanings. We highlight a few examples here (Exhibit 3). However, the purpose of our article is to offer an insight from an accounting history perspective.

From Exhibit 3A., we can see the shapes of notches for Ten Thousand(万), Thousand (千), Hundred (百) and Ten (十). For measuring unity "1″, Exhibit 3B shows that there are at least three kinds of notches depending on the measuring units. Dan (石), Dou (斗) and Sheng (升) are units for measuring the volume of grains. One Dan (石) is equivalent to 10 Dous (斗) and one Dou (斗) is equivalent to 10 Shengs (升). As Dou and Sheng are sub-units for Dan, there is no need for Ten, Hundred, Thousand and Ten Thousand for them. Jin (斤) and Liang (两) are units for measuring weights. One Jin (斤) is equivalent to 16 Liangs (两). Zhang (丈), Chi (尺) and Cun (寸) are units for measuring the length of fabrics. One Zhang (丈) is equivalent to 10 Chis (尺) and one Chi (尺) is equivalent to 10 Cuns (寸). Again, as Liang is a sub-unit for Jin and Chi and Cun are sub-units for Zhang, there is no need for Ten, Hundred, Thousand and Ten Thousand for them.

These slips can be deemed as a form of "dual-method" bookkeeping that combines written details of transactions with corresponding numerical symbols – carved notches. That is, notching repeats a part of bookkeeping in written language on the same slip. In this context, notching is definitely a means of bookkeeping/



(Illustration by James Li)

Exhibit 1. Slips organized with ropes.

accounting. The implication of this is important. While archeologists have speculated that the Middle East clay tokens in Mesopotamia or many other even earlier engraved or etched patterns, such as Blombos Cave engraving, Ishango bone, etc., before the development of written languages, represent ancient bookkeeping/accounting, they could in fact be ancient ornaments or artifacts. We argue that, based on these notched slips, notching/tokening as bookkeeping/accounting can be a valid conjecture.

3. Why did notching reappear during 217–210 BCE?

Bookkeeping slips purely in written language were discovered before this batch of notched slips unearthed in Liye Town. Why did notching reappear? This reappearance could be explained in the historical backdrop of China at that time. The notched slips were made in 217–210 BCE, during China's Qin Dynasty (221 to 207 BCE) when China first became a unified state. Before the Qin Dynasty, China was divided into seven states, Qi, Chu, Yan, Han, Zhao, Wei and Qin, during the Warring States era (475 to 221 BCE). Qin defeated the other six states and unified China. Before unification, each state had its own language. The First Emperor of China (the Qin Emperor) made dramatic reforms in the newly unified China. These reforms included the unification and standardization of the language as well as money, law, measures and the purge of non-Qin schools of thought including Confucianism (Sima, circa 91 BCE a), thus establishing a centralized state. The slips that we examine here were unearthed in Hunan province, which formerly belonged to the State of

#8-817 #8-822 #8-839 #8-903 #8-921 #8-1552 #8-1585 #8-1760

(Image provided by the Research Institute of Cultural Relics and Archaeology of Hunan Province)

Exhibit 2. Notched slips.

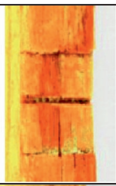
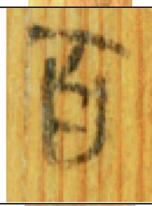
Chu (楚). In fact, archeologists have found official local government language conversion tablets from that time. The process of converging Chu's language to Qin's language caused a temporary dip in literacy in a language that was difficult to learn to start with.

While Qin's central government advocated enhanced literacy for government officials and established official learning institutions (Sima, circa 91 BCE b), Chu was a state famous for its mass resentment against Qin's culture (Sima, circa 91 BCE c) and many refused to abandon the Chu language. That is, local Chu officials who could formerly handle Chu characters were now generally relatively unfamiliar with Qin characters.

With a dip in literacy among government officials in the Chu region, bookkeeping fraud could be more easily committed and more difficult to detect by bookkeeping officials. Note that Qin also unified laws and its laws were known for being extremely austere. Officials accused of cheating were often fined an exorbitant amount of money or faced the prospect of termination of their positions or even death. Qin also adopted "collective punishment" in its laws where related officials would be punished together when one was found to have committed wrongdoing. Therefore, local Chu officials' (those who handled bookkeeping) risk of being accused of cheating increased. Consequently, an increase in the need to prevent bookkeeping fraud coupled with a dip in local literacy in the unified language potentially caused bookkeeping officials to add notching (of numbers and units of measurement) to slips, making an element of the records more indelible.

It is tempting to think that notching appeared to merely temporarily ease the need for people to learn numerical characters or units of measurement. For example, notching required people to know the shapes of "one", "ten", "hundred" and "thousand" in notches, it spared people from learning "two" to "nine", "twenty" and "thirty", some of which would be complicated characters. However, the appearance of notching went beyond just to ease the temporary need for learning numerals or units of measurement. As argued earlier, notching reappeared to satisfy the central government's need to prevent fraud and local officials' desire to avoid being wrongfully accused of bookkeeping fraud, at a time when a dip in local literacy could potentially

Panel A. Notches for ten thousand (万), thousand (千), hundred (百) and ten (十)

Notch	Chinese Character	Abstract Symbol	Number	Slip Serial #
			Ten Thousand（万）	#8-817
			Thousand（千）	#8-839
			Hundred（百）	#8-839
			Ten（十）	#8-1552

Panel B. Notches for one

For measuring unity “1”, there are at least three kinds of notches depending on the measuring units. Dan (石), Dou (斗) and Sheng (升) are units for measuring the volume of grains. One Dan (石) is equivalent to 10 Dous (斗) and one Dou (斗) is equivalent to 10 Shengs (升). Jin (斤) and Liang (两) are units for measuring weights. One Jin (斤) is equivalent to 16 Liangs (两). Zhang (丈), Chi (尺) and Cun (寸) are units for measuring the length of fabrics. One Zhang (丈) is equivalent to 10 Chis (尺) and one Chi (尺) is equivalent to 10 Cuns (寸).

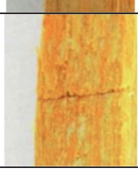
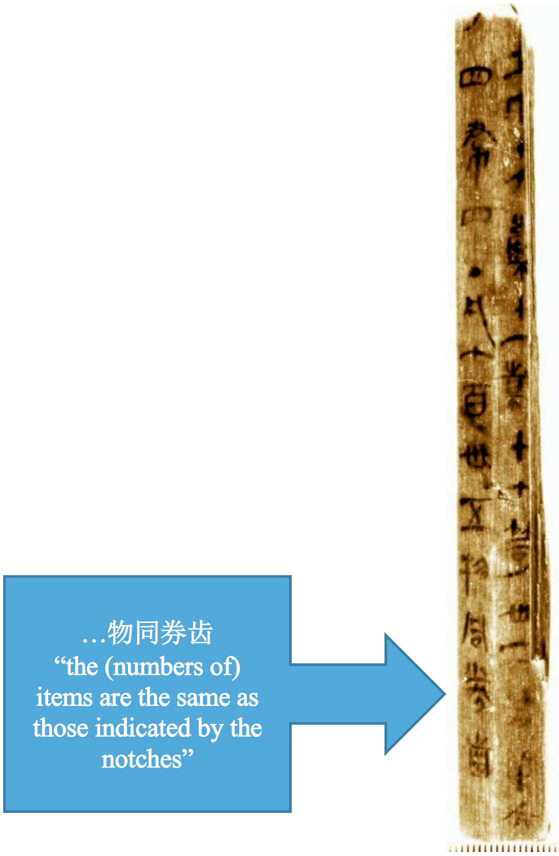
Notch	Chinese Character	Abstract Symbol	Measuring Unit	Slip Serial #
			Dan (石) – Volume of Grain Jin (斤) – Weight Zhang (丈) – Length of Fabrics	#8-822 #8-921 #8-1760
			Dou (斗) – Volume of Grain Liang (两) – Weight Chi (尺) – Length of Fabrics	#8-822 #8-921
			Sheng (升) – Volume of Grain Cun (寸) – Length of Fabrics	#8-1585 #8-1760

Exhibit 3. Meanings of notches based on patterns discussed in Zhang et al. (2015).



【9-2143】☒□巾十六槖十一幭十七笥卅二□十六 ☒□四幚四•凡七百卅五物**同券齿**
(Image provided by the Research Institute of Cultural Relics and Archaeology of Hunan Province)

Exhibit 4. Slip # 9-2143.

make fraud more easily committed and less detectable. In fact, one can imagine an externality of notching being that when the numerical part of a transaction is made more difficult to alter, it reduces the overall desire to falsify other parts of the transaction. Therefore, when language (literacy) failed, notching reappeared, during that time.

Overall, notching provides a means of verification and fraud prevention. In fact, bookkeeping/accounting is a language that requires permanency – the maintenance of permanent records. Notching, even if just the numerical part, makes written records less likely to be altered and thus more permanent. In Slip # 9-2143 (Exhibit 4.) which records textile goods, we see the phrase "…物同券齿" which means "*the* (*numbers of*) *items are the same as those indicated by the notches*". This slip emphasizes the numerical correspondence between the written record and notches. In our modern accounting language, we can consider notching as leaving an "audit trail" in the record-keeping process, an important aspect of bookkeeping/accounting. To this end, the Middle East clay tokens in Mesopotamia or many other even earlier engraved or etched patterns have certainly maintained their permanence.

4. Conclusion

We show a form of "dual-method" bookkeeping that combines written languages as well as carved notches in unearthed wood slips from China during 217–210 BCE. While purely written slips were unearthed before, this batch of slips contain written records of details of transactions as well as notches on their sides that represent the quantities of money or measures of commodities involved. Scholars have speculated that ancient engraving, notching or tokening before the development of written language could be bookkeeping/accounting. We argue that a form of bookkeeping combining notching with written language emerged at a point in Chinese history where a region saw a dip in local literacy and an increase in the austerity of laws. Intriguingly, our analysis of the notched slips supports Childe's (1936) statement that "*Yet the paleolithic sculptures and drawings are not merely expressions of a mysterious 'artistic impulse'. The artist, indeed, surely enjoyed executing them, but he did not do it just to secure that joy, but for a serious economic motive.*" The discovery of these notched slips unearthed in Liye Twon adds credence to the conjecture that prehistorical, pre-language carving, engraving, tokening or notching was likely a form of bookkeeping/accounting. Our analysis adds to our understanding of the origin and evolution of bookkeeping/accounting.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Chen acknowledges funding from the National Social Science Foundation of China: "Research on Historical Materials of Accounting Books in Unearthed Bamboo and Wooden Slips" (Project Approval No.: 22FJYB056).

References

Basu, S., Waymire, G.B., 2006. Recordkeeping and human evolution. Account. Horizon 20, 201–229.

Childe, V.G., 1936. Man Makes Himself. The New American Library.

Henshilwood, C.S., d'Errico, F., Yates, R., Jacobs, Z., Tribolo, C., Duller, G.A.T., Mercier, N., Sealy, J.C., Valladas, H., Watts, I., Wintle, A.G., 2002. Emergence of modern human behavior: Middle Stone Age engravings from South Africa. *Science* 295, 1278–1280.

Henshilwood, C.S., d'Errico, F., van Niekerk, K.L., Dayet, L., Queffelec, A., Pollarolo, L., 2018. An abstract drawing from the 73,000-year-old levels at Blombos Cave, South Africa. *Nature* 562, 115–118.

Hovers, E., 1990. Art in the Levantine Epi-Palaeolithic: an engraved pebble from a Kebaran site in the lower Jordan valley. Curr. Anthropol. 31 (3), 317–322.

Mackay, A., Welz, A., 2008. Engraved ochre from a Middle Stone Age context at Klein Kliphuis in the Western Cape of South Africa. *J. Archaeol. Sci.* 35, 1521–1532.

Marshack, A., 1979. Upper paleolithic symbol systems of the Russian plain: cognitive and comparative analysis. *Curr. Anthrop.* 20, 271–311.

Marshack, A., 1996. A middle Paleolithic symbolic composition from the Golan Heights: the earliest known depictive image. *Curr. Anthropol.* 37 (2), 357–365.

Mattessich, R., 1987. Prehistorical accounting and the problem of representation: on recent archeological evidence of the Middle-east from 8,000 B.C. to 3000 B.C. *Account. Hist. J.* 14, 71–91.

Mattessich, R., 1994. Archaeology of accounting and Schmandt-Besserat's contribution. *Account. Bus. Financ. Hist* 4, 5–28.

Peng, F., Gao, X., Wang, H., Chen, F., Liu, D., Pei, S., 2012. The discovery of scratched stoneware of the late Paleolithic age at the Shuidonggou site (水洞沟旧石器时代晚期遗址发现带有刻划痕迹的石制品). Chin. Sci. Bull. 57 (26), 2475–2481.

Schmandt-Besserat, D., 1983. Token and counting. *Biblical Archaeol.* 46, 117–120.

Sima, Q., (司马迁). circa 91 BCE a. The Grand Scribe's Records – The First Emperor of Qin (《史记·秦始皇本纪》).

Sima, Q., (司马迁). circa 91 BCE b. The Grand Scribe's Records – Li Si (《史记·李斯列传》).

Sima, Q., (司马迁). circa 91 BCE c. The Grand Scribe's Records – Xiang Yu (《史记·项羽本纪》).

Zhang, C., Toshitaka, O., Akira, M., 2015. A research on notched Liye slips (里耶秦简刻齿简研究 – 兼论岳麓秦简《数》中的未解读简). *Cult. Relics* 3, 53–96.

China Journal of Accounting Research 18 (2025) 100434

HOSTED BY

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Annual report audit, ESG report assurance and audit quality: Evidence from the same accounting firm

Wenwen Li [a], Ting Li [b,*], Hongjun Zhu [c]

[a] School of Accounting, Shanghai University of Finance and Economics, China
[b] Institute of Accounting and Finance, Shanghai University of Finance and Economics, China
[c] School of Accounting, Anhui University of Finance and Economics, China

ARTICLE INFO

Article history:
Received 13 May 2024
Accepted 19 June 2025
Available online 5 July 2025

Keywords:
ESG
Independent assurance
Audit quality
Knowledge spillover
Reputation investment

ABSTRACT

This study investigates the impact of an accounting firm's providing both an annual report audit and environmental, social and governance (ESG) report assurance on its audit quality. We find that when the same accounting firm is responsible for both the annual report audit and ESG report assurance, this does not compromise the accounting firm's independence. On the contrary, it enhances audit quality through knowledge spillover effects and increased investment in reputational capital. Additional evidence suggests that providing ESG report assurance also prompts the accounting firm to allocate more audit resources, thereby influencing audit quality. In cross-sectional results, consistent with expectations, we find that the knowledge spillover and reputational effects of ESG report assurance are more pronounced for companies with weaker internal controls, non-A + H share companies and companies receiving lower levels of analyst attention. Finally, we document that having the same accounting firm responsible for ESG report assurance and annual report audit increases audit fees and also contributes to enhanced firm value. This study comprehensively highlights the influence of ESG report assurance on audit quality, and its findings offer valuable insights into the economic consequences of ESG report assurance practices.

1. Introduction

In recent years, an increasing number of stakeholders, including investors, consumers and regulatory authorities, have begun incorporating firms' environmental, social and governance (ESG) performance into their

* Corresponding author.
E-mail addresses: lww@163.sufe.edu.cn (W. Li), li.ting@mail.shufe.edu.cn (T. Li), 120200088@aufe.edu.cn (H. Zhu).

https://doi.org/10.1016/j.cjar.2025.100434
1755-3091/

decision-making frameworks (Hartzmark and Sussman, 2019; Shi et al., 2024). Accurate and effective ESG disclosure, as a crucial source of information on corporate ESG performance, not only enhances transparency but also provides stakeholders with the necessary basis to evaluate a firm's long-term sustainability and risk management, thereby underscoring the enhanced significance of ESG reports. However, as ESG disclosure is still in its nascent stages, the standards and reporting frameworks for corporate ESG disclosures remain incompletely unified. ESG information is relatively unstandardized and lacking quantitative detail, resulting in a generally poor quality of disclosure (Shen et al., 2023), with some firms even engaging in "greenwashing" (Huang, 2022; Hu et al., 2023). ESG reports require independent third-party assurance to ensure their quality and to address the prevalent greenwashing phenomenon. The independence of assurance providers directly affects the objectivity and reliability of ESG reports and annual reports. A pertinent issue is whether providing both annual report audit and ESG report assurance services simultaneously increases an accounting firm's economic dependence on clients (Frankel et al., 2002) and consequently impairs auditor independence. Previous studies find that ESG reports contain significant informational value that can trigger notable market reactions (Krueger et al., 2024). As it primarily focuses on non-financial information, ESG assurance may provide insights and information that benefit the audit work (Simunic, 1984). For instance, specific information on a firm's green credit discerned through ESG assurance can be compared with the credit information in the annual report, aiding in identifying potential risks related to environmental responsibility and governance structures, and potentially facilitating the auditor's annual report audit work. Little prior research discusses issues related to ESG assurance, and even fewer studies directly explore whether simultaneously undertaking ESG report assurance and annual report audit services impacts accounting firms. In this study, we help close this gap by examining whether accounting firms' provision of both annual report audit and ESG assurance services simultaneously impairs auditor independence and thus diminishes audit quality, or if such accounting firms leverage knowledge spillover and reputational effects to benefit both themselves and their clients.

As the world's second largest economy, China's market size and development potential exert a profound influence on the global economy. With increasing attention from international investors to corporate sustainability and corporate social responsibility (CSR), exploring the practices of Chinese firms in ESG information disclosure and assurance can provide valuable insights for global markets (Shen et al., 2023). At the same time, the Chinese government has adopted proactive policy measures to promote sustainable development, continuously issuing policy documents related to corporate ESG information disclosure (Huang, 2022). From the perspective of ESG information disclosure assurance among Chinese listed companies, a limited number of firms that have implemented such disclosures have already engaged in assurance procedures. Many of these companies have hired the same accounting firms responsible for their annual report audit to provide assurance services for their ESG report as well. This unique setting presents an opportunity to examine whether the provision of both ESG assurance and annual report audit services by the same accounting firm affects auditor independence. Therefore, we test our hypothesis in the Chinese setting, aiming to provide evidence relevant to ESG research and practice worldwide.

By studying cases in which the same accounting firm provides both ESG report assurance and annual report audit services, this empirical study makes the following observations. (1) When the same accounting firm is responsible for both annual report audit and ESG report assurance, such an arrangement does not impair the accounting firm's independence. On the contrary, this collaboration facilitates knowledge spillover and enhances the accounting firm's professional competence through reputational capital investments, thereby improving audit quality. (2) Mechanism analysis reveals that undertaking annual report audit and ESG assurance services concurrently encourages accounting firms to increase their resource allocation to audit engagements, which positively impacts audit quality. (3) Further analysis shows that the impact of simultaneously providing annual report audit and ESG assurance services on audit quality is more pronounced for firms with weaker internal controls, non-A + H share firms and firms with lower analyst attention. (4) Economic consequence tests indicate that accounting firms responsible for ESG assurance services charge higher audit fees, while also contributing to an enhancement in the firm's value. Given that the sample in this study only includes listed companies with ESG report assurance services, to address potential endogeneity issues such as sample selection bias and reverse causality, we conduct several robustness tests, including focusing on a propensity score matched sample, employing the Heckman model and using instrumental variable (IV) approaches. Additionally, we rule out alternative explanations for the results by comparing subsamples with varying characteristics.

Our analysis makes several contributions. First, prior research predominantly focuses on the impact of accounting firms' providing non-audit services (such as management consulting or tax planning) on audit quality (Defond et al., 2002; Ao and Sun, 2024). Non-audit services such as management or tax consulting primarily involve financial management, tax planning and operational optimization, targeting internal management with strong direct ties to financial reports but offering limited contributions to assessing a firm's long-term sustainability. In contrast, ESG report assurance, while part of audit services, focuses on non-financial information (i.e., environmental, social and governance issues), aiming to enhance the credibility of ESG disclosures and improve the completeness of financial reports. ESG information can be used to identify environmental responsibilities and governance risks in financial reports, thereby producing broader knowledge spillover effects. Additionally, ESG report assurance can bolster the market reputation of firms, enhancing investor trust in audit outcomes without excessively compromising audit independence, as management or tax consulting might. Thus, investigating how accounting firms' providing ESG assurance services affects audit quality from the perspective of non-financial report audit services distinguishes this study from traditional non-audit service research. Furthermore, previous studies on the impact of accounting firms' providing audit services on audit quality primarily explore how offering internal control audit services affects the quality of financial report audit from an integrated audit perspective (Ni and Zhang, 2015; Zheng et al., 2015). Few studies investigate whether auditors' simultaneous responsibility for annual report audit and ESG report assurance impacts their independence, and this research fills that gap from an audit quality perspective. Internal control audits focus on financial systems and internal governance, while ESG report assurance involves non-financial information. This study extends existing research, which mainly emphasizes financial information audit services, by examining the impact of accounting firms' providing non-financial information audit services on audit quality.

Second, prior research on ESG report assurance mainly examines the factors influencing ESG report assurance and whether ESG reports contain informational value (Li and Li, 2012; Li et al., 2013; Krueger et al., 2024), with limited studies discussing how ESG report assurance affects auditor decision-making. Similar to this study, using data from 55 countries and regions, Maso et al. (2020) find that accounting firms providing CSR assurance services for the same client, compared with those only offering annual report audit, more frequently issue going concern opinions. However, Maso et al.'s (2020) data are affected by issues such as varying collection methods, standards, frequencies and quality, potentially affecting the credibility of cross-national comparisons. Moreover, considering the significant differences in institutional, cultural and religious backgrounds among countries, cross-national studies might overlook certain social factors, possibly leading to biased outcomes. Consequently, accurately understanding how providing ESG report assurance affects audit quality requires more precise empirical data. This study focuses on the Chinese context, using only data from China, which effectively eliminates the influence of national-level characteristics on the research conclusions. Thus, the findings of this study enhance the credibility and scientific rigor of Maso et al. (2020) and serve as a substantial addition to the literature.

Third, this study concludes that accounting firms' provision of both annual report audit and ESG report assurance services does not impair audit independence. Instead, it improves audit quality through knowledge spillover and increased investment in reputational capital, providing theoretical support for the positive effects of accounting firms' undertaking ESG report assurance services. This holds significant practical implications.

The rest of the paper is organized as follows. Section 2 provides the institutional background and literature review; Section 3 presents the theoretical analysis and develops our hypotheses; Section 4 outlines the research design; Section 5 reports the main regression results; Section 6 offers further analysis and examines the economic consequences; and Section 7 draws a conclusion.

2. Institutional background and literature review

2.1. Institutional background

Currently, regulatory authorities in mainland China rarely mandate the preparation and disclosure of ESG reports by listed companies, nor do they clearly require assurance for ESG reports. In contrast, the Hong Kong Stock Exchange issued its first "Environmental, Social and Governance Report Guide" in 2012, recommending that listed companies disclose key ESG indicators in their annual reports or standalone reports, marking a "voluntary disclosure" phase. In 2015, the Stock Exchange issued the revised "ESG Report Guide (2015 version)" and introduced a "comply or explain" principle, transitioning ESG report requirements to a semi-mandatory phase. On 18 December 2019, the Stock Exchange officially released the newly revised "ESG Guide", which requires listed companies to publish ESG reports in accordance with the updated guidelines, thus moving ESG disclosure requirements into a mandatory phase. However, the requirement for assurance of ESG reports remains non-mandatory; companies are merely encouraged to undertake independent assurance of their ESG disclosures. Observing the evolution of the ESG reporting and assurance requirements of Hong Kong's regulatory authority, it is anticipated that in the future, listed companies in mainland China will also need to prepare and disclose ESG reports, along with obtaining assurance for these reports. The primary goal of ESG report assurance is to validate, through independent auditing procedures, whether the disclosed ESG information meets relevant standards and to ensure the credibility of the data, thereby better supporting investment decisions.

The main aspects involved in ESG assurance include the assurance content, standards and methods and the level of assurance. As set out in the "T/CERDS 8-2024 Enterprise ESG Report Assurance Guide", the assurance content of ESG reports mainly covers the three domains of environmental (E), social (S) and governance (G).[1] The standards for ESG report assurance are primarily used to assess whether the ESG information disclosed by enterprises meets relevant requirements, ensuring that the data are truthful, accurate and complete, thereby providing stakeholders with a reliable basis for decision-making. ESG report assurance standards encompass both global international standards and local Chinese standards, applicable to various scenarios of ESG report assurance. The level of ESG assurance, which determines the depth of the assurance work, is mainly classified into reasonable and limited assurance. Reasonable assurance involves greater depth, requiring more time and stricter audit procedures, while limited assurance is a lighter procedure, more suitable for ESG reports not requiring in-depth examination, and is appropriate for enterprises in the early stages of information disclosure. Panels A and B of Table 1 list the ESG report assurance standards and levels of assurance, respectively.

2.2. Literature review

Prior studies mostly discuss the factors influencing audit quality from the perspectives of institutional frameworks and individual auditors. From an institutional standpoint, the stricter the regulation and the higher the legal risk, the higher the audit quality (Khurana and Raman, 2004; Gipper et al., 2020). At the individual auditor level, industry expert auditors disclose a greater number of key audit matters, produce longer report texts and are more inclined to adopt conclusive evaluations than non-experts (Stein, 2019). Additionally, the sharing of personal industry knowledge among audit partners within a firm can reduce the frequency of financial restatements and enhance audit quality (Michas et al., 2025). An auditor's personal experience significantly contributes to audit work, as experienced auditors possess more knowledge about financial reporting misstatements, thereby increasing the likelihood of detecting errors in financial reports (Libby and Frederick, 1990). Consequently, the more experienced the auditor, the higher the audit quality (Cahan and Sun, 2015). Regarding the impact of auditor gender, female auditors exhibit greater caution and risk aversion than their male counterparts, leading to lower levels of negative earnings management in financial reports audited by female auditors (Yang et al., 2018).

Theoretical research on the impact of auditors' providing non-audit services on those auditors' audit services generally takes either the beneficial view or the detrimental view. The beneficial view argues that auditors' provision of non-audit services can generate knowledge spillover effects—a concept first introduced by Simunic (1984)—thereby improving audit quality by leveraging the knowledge gained from non-audit services. In contrast, the detrimental view is rooted in fee dependence theory, which posits that the inclusion of non-audit services strengthens the client's bargaining power over the auditor. As a result, providing both audit and non-audit services may compromise the independence of the audit firm, ultimately harming audit quality

[1] For specific examples of ESG report assurance content, see Appendix A.

Table 1
Panel A: ESG report assurance standards.

Standard type	Standard name	Scope of applicability	Main content
International standards	ISAE 3000	Assurance of ESG information	Used for independent assurance of non-financial information, ensuring data authenticity and completeness.
	ISAE 3410	Greenhouse gas emissions assurance	Specifically used for carbon emission verification, ensuring accuracy of greenhouse gas data.
	AA1000AS v3	Sustainability report assurance	Focuses on the credibility of corporate ESG strategies and information disclosure.
	ISO 14064-3	Carbon emissions verification	Applicable for independent review of greenhouse gas emission data.
Chinese standards	T/CERDS 8-2024	ESG report assurance	Provides a comprehensive process for enterprise ESG report assurance.
	GB/T 24353	Risk management	Used for risk assessment within ESG assurance.
	T/CERDS2	ESG disclosure guidelines	Regulates the content of corporate ESG information disclosure.
	GB/T 32150 & GB/T 32151	Carbon emissions accounting	Applicable for calculation and verification of corporate carbon footprints.

Panel B: Levels of ESG report assurance.

Level of assurance	Characteristics	Applicable scenarios	Assurance methods	Assurance conclusion
Reasonable assurance	The highest level of assurance, similar to financial auditing, requiring extensive data verification, on-site investigations and analysis.	Publicly listed companies under stringent regulatory requirements, or enterprises where ESG information significantly impacts investment decisions.	In-depth data analysis, detailed document review and cross-verification, on-site inspections and stakeholder interviews.	Positive expression of conclusions, such as “In all material respects, the content of this ESG report is reliable”.
Limited assurance	A lower level of assurance than reasonable assurance, primarily based on data provided by the company and analyzed using sampling methods.	Smaller-scale enterprises or companies with limited ESG disclosure scope.	Relies mainly on data provided by the company, interviews with selected key personnel and limited cross-verification.	Negative expression of conclusions, such as “No material misstatements or omissions have been identified”.

(Frankel et al., 2002). Several studies indicate that providing non-audit services does not undermine auditor independence; on the contrary, it may enhance auditors’ investment in reputational capital, reduce audit errors and lower the probability of misreporting (Dopuch et al., 2001). The synergies and knowledge spillover effects generated by non-audit services can also contribute to improved audit quality (Simunic, 1984). In China, research on this topic mainly adopts an integrated audit perspective, exploring the impact of auditors’ simultaneously providing internal control audit and annual report audit services. Fang and Chen (2016) find a significant positive relationship between internal control audit fees and annual report audit fees, with evidence supporting the claim of Ni and Zhang (2015) that integrated audits can improve audit quality, consistent with knowledge spillover theory. However, Yang et al. (2017) find that compared with non-integrated audits, integrated audits generate scale synergies, improve audit transaction efficiency and reduce audit fees, while not finding significant evidence of differences in audit quality under different audit models.

Research on the economic consequences of ESG report assurance is relatively scarce. As mentioned above, using data from over 50 countries and regions, Maso et al. (2020) find that accounting firms providing both annual report audit and CSR assurance services for the same client are more likely to issue going concern opinions. Li and Li (2012) and Li et al. (2013) focus on the factors influencing corporate ESG report assurance activities and whether ESG report assurance opinions carry informational value, showing that ESG report assurance opinions generate positive market reactions. Moreover, publicly listed companies exposed in the media for negative CSR incidents are less likely to engage in ESG report assurance, while companies located in regions with stronger legal systems and higher trust levels are more inclined to obtain ESG report assurance. From the above findings, it is evident that there is still a lack of research investigating whether auditors’ simultaneously providing ESG report assurance and annual report audit services affects their decision-making. Therefore, this paper explores whether simultaneously offering ESG report assurance services impairs the independence of accounting firms or, alternatively, improves audit quality through knowledge spillover effects and increased investment in reputational capital.

3. Hypothesis development

Auditors that provide both annual report audit and ESG report assurance services can offer enhanced audit quality through a knowledge spillover effect (Simunic, 1984; Zheng et al., 2015). Simunic (1984) introduced the concept of knowledge spillover in the context of the impact of auditors' providing non-audit services on their audit quality. This concept suggests that auditors can acquire additional knowledge and information when delivering non-audit services, which may "spill over" into their audit work. Such spillover can enhance auditors' expertise, thereby improving the overall audit quality.

The ESG report assurance service is primarily a process by which the assurer conducts a scientific audit of a company's non-financial information and expresses an opinion. From the perspective of the knowledge spillover effect, auditors can acquire non-financial information, beyond that included in the annual report, through ESG report assurance, thereby assisting the annual report audit work. For instance, by assuring ESG reports, auditors can obtain information related to environmental decommissioning costs or risks associated with debt repayment. Specifically, KPMG was responsible for both the 2021 annual report audit and ESG report assurance for Shanghai Pudong Development Bank Co., Ltd. Unlike the 2021 annual report, the bank's 2021 ESG report provided detailed listings of its green credit projects and related amounts. While conducting the assurance of the ESG report, KPMG performed analytical procedures on these key data and cross-verified them with the relevant accounting records to ensure data consistency. Similarly, China Shenhua Energy Company Limited did not disclose information related to coal safety production in its 2021 annual report but explicitly advised investors to refer to its 2021 ESG report for details on coal safety production. KPMG, which was responsible for both the annual report audit and the ESG report assurance for this firm, reviewed key indicators related to safety production during the ESG report assurance process. This illustrates how auditors can gain additional information from ESG reports (Shi et al., 2024), which aids in conducting substantive procedures and corresponding tests, helping auditors identify potential environmental risks and deficiencies in internal control management disclosed in the annual report, thereby enhancing their understanding of the client. This spillover of knowledge and information enhances the quality of an auditor's annual report audit. In addition, the spillover of knowledge and information enhances the auditor's professional competence. The stronger auditors' professional ability, the more likely they are to detect manipulations in companies' annual reports, thus improving audit quality (Fan et al., 2013).

From a reputation perspective, providing ESG report assurance services encourages accounting firms to invest more in reputational capital, thereby reducing the likelihood of auditors' succumbing to pressure from client firms. To secure more ESG report assurance engagements and expand their market share in the ESG report assurance sector, accounting firms must maintain a strong reputation. Increased investment in reputational capital does not impair auditor independence but instead enhances it, which is beneficial for reputation collateral effects. This helps reduce audit bias, lowers the probability of misreporting and ultimately improves audit quality (Dopuch et al., 2001; Zheng et al., 2015). Moreover, ESG report assurance services are currently primarily provided by large, highly reputable accounting firms, which face higher costs for compromising their independence due to client pressures. Additionally, based on the "deep pockets" idea, such accounting firms are exposed to higher litigation costs in the event of audit failures, further incentivizing them to improve audit quality (Simunic and Stein, 1996).

In summary, through the knowledge spillover effect, providing ESG report assurance services alongside financial audits encourages auditors to engage in more substantive procedures, such as inventory counts, confirmations and risk testing, to investigate any discrepancies between the ESG report and the annual report. Additionally, the reputational benefits of investment in reputational capital depend on increased audit efforts to avoid audit failures. Therefore, this paper posits that accounting firms' offering of both annual report audit and ESG report assurance services can enhance audit investment through the knowledge spillover and reputation effects, thereby improving audit quality. This rationale underlies our prediction as follows.

H1: The provision of both ESG report assurance and annual report audit services by the same accounting firm enhances audit quality.

However, concurrently providing ESG report assurance and annual report audit services increases the level of economic contracting between accounting firms and their clients, which may exacerbate auditors' economic dependence on their clients (Beck et al., 1998; Beeler and Hunton, 2001). This can undermine auditor independence, potentially causing auditors to compromise audit quality due to client pressure. As audit quality is contingent on the professionalism and independence of accounting firms, offering ESG report assurance services may enhance auditors' professionalism through knowledge spillover and reputation effects, thereby improving audit quality. Nevertheless, this practice may also heighten the accounting firm's economic reliance on the client company, diminishing independence and consequently reducing audit quality. Therefore, this issue warrants further empirical examination to explore its implications fully.

4. Research design

4.1. Sample selection and data sources

We use Chinese A-share listed companies from 2009 to 2023 as our research sample. Due to the limited availability of ESG report assurance data prior to 2009, we start our sample period in that year. We are interested in whether accounting firms' provision of both annual report audit and ESG report assurance services impacts their audit quality, compared with other accounting firms or assurance providers. Thus, the sample exclusively includes listed companies with ESG report assurance. Additionally, the following data processing steps are taken: (1) exclusion of companies under special treatment; and (2) exclusion of companies with missing financial and corporate governance data. This results in a final dataset of 595 observations. Data on whether the same accounting firm provides both ESG report assurance and annual report audit services are manually matched (specific processing details are provided in Section 4.2). Other data are sourced from the China Stock Market & Accounting Research database. To mitigate the impact of extreme values, we winsorize all continuous variables at both the 1st and 99th percentiles.

4.2. Variable definitions

4.2.1. Audit quality (Restate)

Drawing on Pittman et al. (2023) and Dou et al. (2024), we measure audit quality by the occurrence of financial restatements in annual reports. Specifically, if a company's annual report contains significant accounting errors that result in a restatement in subsequent years, the variable *Restate* is assigned a value of one; otherwise, its value is zero.

4.2.2. Whether ESG report assurance and annual report audit services are provided by the same accounting firm (Same)

We retain only those companies whose ESG reports have undergone assurance. The provision of ESG report assurance services is determined based on the name of the institution providing the assurance; if the name corresponds to an accounting firm, further verification is performed to determine whether it is consistent with the name of the accounting firm responsible for auditing the annual report. If the ESG report assurance and the annual report audit are provided by the same accounting firm (with matching names) and are located in the same area, it is considered that both the ESG report assurance and annual report audit are provided by the same accounting firm. In such cases, the variable *Same* is assigned a value of one; otherwise, it is set to zero.[2]

[2] For example, for Hua Xia Bank Co., Limited, the institution providing assurance for its 2021 ESG report was Ernst & Young (LLP) (Beijing), and the auditor for its 2021 annual report was also Ernst & Young (LLP) (Beijing). In this case, it is considered that both the ESG report assurance and the annual report audit are provided by the same accounting firm, and the variable *Same* is assigned a value of one. Conversely, in 2012, the assurance for its ESG report was conducted by Bureau Veritas Certification, while the auditor for the annual report was Grant Thornton International Ltd (LLP). Therefore, it is concluded that the ESG report assurance and the annual report audit are not provided by the same accounting firm, and *Same* is assigned a value of zero.

4.2.3. Audit input (Delay)

Following Zhang (2018), we use audit delay as a measure of audit effort. *Delay* is calculated as the number of days between the financial statement date and the audit report release date. A longer audit delay implies a greater level of audit effort.

4.2.4. Other control variables

Drawing on existing research, we select the following factors as control variables: firm size (*Size*), leverage ratio (*Lev*), return on assets (*Roa*), loss status (*Loss*), proportion of independent directors (*Ddratio*), board size (*Dshsize*), firm age (*Age*), auditor choice (*Big4*), state-ownership status (*Soe*), dual listing status in Hong Kong (*AH*), book-to-market ratio (*BM*) and audit tenure (*Tenure*). We provide detailed variable definitions in Table 2.

4.3. Model

As the dependent variable, *Restate*, is a binary variable, we employ a logit model for testing while controlling for industry and year fixed effects. In accordance with the hypotheses proposed in this paper, the following model is constructed:

$$\begin{aligned} Restate = {} & \alpha_0 + \alpha_1 Same + \alpha_2 Size + \alpha_3 Lev + \alpha_4 Roa + \alpha_5 Loss + \alpha_6 Ddratio \\ & + \alpha_7 Dshsize + \alpha_8 Age + \alpha_9 Big4 + \alpha_{10} Soe + \alpha_{11} AH + \alpha_{12} BM + \alpha_{13} Tenure + Year\,Fixed \\ & + Industry\,Fixed + \varepsilon \end{aligned} \tag{1}$$

Table 2
Variable definitions.

Variable type	Variable	Variable definitions
Dependent and independent variables	*Restate*	Audit quality, indicated by whether the annual report of a listed company is restated; coded as one if restated, and zero otherwise.
	Same	Indicates whether ESG report assurance and annual report audit are conducted by the same accounting firm; coded as one if yes, and zero otherwise.
	Delay	Audit effort, measured by the number of days between the financial statement date and the audit report release date.
	Fee	Audit fees, quantified as the natural logarithm of the audit fees disclosed in the annual report of the listed company.
	Tbinq	Firm value, defined as the firm's market value of assets divided by book value of assets.
Control variables	*Size*	Firm size, measured as the natural logarithm of the firm's total assets.
	Lev	Leverage ratio, calculated as total liabilities divided by total assets.
	Roa	Return on assets, calculated as net profit divided by total assets.
	Loss	Loss status, indicated as one if the firm incurs a loss, and zero otherwise.
	Ddratio	Proportion of independent directors, calculated as the number of independent directors divided by the total number of board members.
	Dshsize	Board size, measured by the total number of board members.
	Age	Firm age, measured as the natural logarithm of the number of years since listing.
	Big4	A dummy variable equal to one if a firm is audited by one of the international Big 4 auditors, and zero otherwise.
	Soe	State-ownership dummy, defined as a dummy variable equal to one if the firm is a state-owned enterprise, and zero otherwise.
	AH	Indicates whether the firm is also listed on the Hong Kong Stock Exchange; coded as one if yes, and zero otherwise.
	BM	Book-to-market ratio, calculated as total assets divided by the firm's market value of assets.
	Tenure	Audit tenure, measured by the number of consecutive years for which the accounting firm has audited the listed firm.

Based on the previous analysis, when an accounting firm provides both annual report audit and ESG report assurance services, this is expected to enhance audit quality through the knowledge spillover effect and reputational effect. The coefficient of α_1 is anticipated to be negative and significant.

To examine the underlying mechanism, we draw on Wen and Ye (2014) and employ a mediating effect model to test whether accounting firms' concurrent undertaking of annual report audit and ESG report assurance services enhances audit quality by increasing audit effort. Models (2) and (3) are constructed for this purpose:

$$\begin{aligned} Delay = {} & \alpha_0 + \alpha_1 Same + \alpha_2 Size + \alpha_3 Lev + \alpha_4 Roa + \alpha_5 Loss + \alpha_6 Ddratio + \alpha_7 Dshsize + \alpha_8 Age + \alpha_9 Big4 \\ & + \alpha_{10} Soe + \alpha_{11} AH + \alpha_{12} BM + \alpha_{13} Tenure + Year\,Fixed + Industry\,Fixed + \varepsilon \end{aligned} \tag{2}$$

$$\begin{aligned} Restate = {} & \alpha_0 + \alpha_1 Same + \alpha_2 Delay + \alpha_3 Size + \alpha_4 Lev + \alpha_5 Roa + \alpha_6 Loss + \alpha_7 Ddratio \\ & + \alpha_8 Dshsize + \alpha_9 Age + \alpha_{10} Big4 + \alpha_{11} Soe + \alpha_{12} AH + \alpha_{13} BM + \alpha_{14} Tenure + Year\,Fixed \\ & + Industry\,Fixed + \varepsilon \end{aligned} \tag{3}$$

Audit effort is measured using *Delay*. Building on regression model (1), model (2) represents the second step of the mediating effect test, where the coefficient of α_1 is expected to be positive and significant. Model (3) constitutes the third step in testing the mediating effect; if the mediating effect is established, the coefficient of α_2 is expected to be negative and significant.

5. Empirical results

5.1. Descriptive statistics

Table 3 presents the descriptive statistics of the data used in this study. Panel A illustrates the annual distribution of ESG report assurance services for listed companies. Overall, 47.73 % of ESG reports are assured by accounting firms, while 34.62 % (29.92 %) of the firms have their annual report audit and ESG report assurances provided by the same accounting firm (the same Big 4 accounting firm). Panel B presents the industry distribution of ESG report assurance services for listed companies. The financial industry has the largest number of assured ESG reports, with 244 observations. Panel C shows the distribution of ESG report assurance services for listed companies by the nature of the firm. Compared with non-state-owned enterprises, A-share-only listed firms and firms audited by non-Big 4 firms, state-owned enterprises, A + H share listed firms and firms audited by the Big 4 are more likely to have their ESG reports assured.

Panel D presents the distribution of types of institutions providing ESG report assurance that are not accounting firms. This primarily includes ESG service providers, standard-setting organizations, accreditation bodies, consulting firms, and third-party testing, inspection and certification entities. Among the sample of companies that opt for ESG report assurance from non-accounting firms, 169 select a Big 4 firm for auditing their annual report. Meanwhile, among the sample of companies that choose other accounting firms for ESG report assurance services, 64 select the Big 4 for auditing their annual report. Panel E outlines the annual information regarding the provision of ESG reports and their assurance by publicly listed companies during the sample period. Approximately 30% of publicly listed companies provide ESG reports, while only 4.26% of them undergo assurance of their ESG report. This finding is largely consistent with the results of Shen et al. (2023).

Panel F reports the summary statistics for the main regression variables. The mean of *Restate* is 0.091, indicating that 9.1% of the listed companies in the sample have financial restatements. The mean of *Same* is 0.346, which suggests that among the companies with assured ESG reports, 34.6% receive assurance services from the same accounting firms responsible for their annual report audit. The mean of *Big4* is 0.687, indicating that 68.7% of the companies in this study are audited by Big 4 accounting firms. The statistical results for other variables are generally consistent with existing research.

Table 3
Descriptive statistics.

Panel A: Annual distribution.

Year	ESG reports assured	Assured by accounting firms	Not assured by accounting firms	Proportion assured by accounting firms (%)	Same accounting firm for annual and ESG reports (Big 4)	Not same accounting firm for annual and ESG reports	Proportion of same accounting firm (Big 4) (%)
2009	10	5	5	50	5	5	50
					(4)		(40)
2010	19	11	8	57.89	10	9	52.63
					(7)		(36.84)
2011	24	13	11	54.17	12	12	50
					(9)		(37.5)
2012	33	17	16	51.52	13	20	39.39
					(10)		(30.30)
2013	26	18	8	69.23	11	15	42.31
					(9)		(34.62)
2014	26	15	11	57.69	11	15	42.31
					(9)		(34.62)
2015	26	16	10	61.54	11	15	42.31
					(9)		(34.62)
2016	28	17	11	60.71	11	17	39.29
					(9)		(32.14)
2017	32	17	15	53.13	11	21	34.38
					(10)		(31.25)
2018	30	18	12	60	13	17	43.33
					(12)		(40)
2019	22	19	3	86.37	16	6	27.27
					(15)		(68.18)
2020	42	21	21	50	17	25	40.48
					(16)		(38.1)
2021	41	21	20	51.22	15	26	36.59
					(14)		(34.15)
2022	89	34	55	38.20	21	68	23.60
					(19)		(21.35)
2023	147	42	105	28.57	29	118	19.73
					(24)		(16.33)
Total	595	284	311	47.73	206	389	34.62
					(178)		(29.92)

Panel B: Industry distribution.

Industry	ESG reports assured	Assured by accounting firms	Not assured by accounting firms	Proportion assured by accounting firms (%)	Same accounting firm for annual and ESG reports	Not same accounting firm for annual and ESG reports	Proportion of same accounting firm (%)
Mining	40	26	14	65	24	16	60
Manufacturing	165	15	150	9.1	6	159	3.6
Production and supply of electricity, heat, gas and water	20	6	14	30	1	19	5
Construction	15	1	14	6.67	0	15	0
Wholesale and retail trade	29	18	11	62.07	18	11	62.07
Transportation, storage and postal services	47	4	43	8.51	2	45	28.57
Accommodation and catering	1	0	1	0	0	1	0
Information transmission, software and information technology services	5	1	4	20	1	4	20
Financial	244	198	46	81.15	148	96	60.66
Real estate	13	4	9	30.77	4	9	30.77
Leasing and business services	6	5	1	83.33	0	6	0
Scientific research and technical services	6	0	6	0	0	6	0
Water conservancy, environment and public facility management	2	1	1	50	1	1	50
Healthcare and social work	1	1	0	100	1	0	100
Comprehensive category	1	1	0	100	0	1	0
Total	595	284	311	47.73	206	389	34.62

Panel C: Distribution of enterprise types.

Enterprise type	ESG reports assured	Assured by accounting firms	Not assured by accounting firms	Proportion assured by accounting firms (%)	Same accounting firm for annual and ESG reports	Not same accounting firm for annual and ESG reports	Proportion of same accounting firm (%)
State-owned enterprises	340	187	153	55	145	195	42.65
Non-state-owned enterprises	255	97	158	38.04	61	194	23.92
A-Share	293	98	195	33.45	63	230	21.50
A+H Share	302	186	116	61.59	143	159	47.35
Big 4	409	240	169	58.68	176	233	43.03
Non-Big 4	186	44	142	23.66	30	156	16.13
Total	595	284	311	47.73	206	389	34.62

Panel D: Types of assurance institutions.

Non-accounting firms providing ESG assurance	Number	Types of accounting firms auditing annual report (non-accounting firms providing ESG report assurance)	Number	Types of accounting firms auditing annual report (other accounting firms providing ESG report assurance)	Number
ESG service providers	12	Big 4	169	Big 4	64
Standard-setting organizations	10	Non-Big 4	142	Non-Big 4	14
Accreditation bodies	13	/	/	/	/
Consulting firms	77	/	/	/	/
Third-party testing, inspection and certification entities	199	/	/	/	/
Total	311	/	311	/	78

Panel E: Distribution of listed companies.

Year	Number of publicly listed companies	Number of companies providing ESG report	Percentage (%)	Number of assured ESG reports	Percentage (%)
2009	1399	173	12.37	10	5.78
2010	1623	447	27.54	19	4.25
2011	1903	549	28.85	24	4.37
2012	2343	637	27.19	33	5.18
2013	2363	668	28.27	26	3.89
2014	2427	682	28.10	26	3.81
2015	2597	715	27.53	26	3.64
2016	2854	767	26.87	28	3.65
2017	3277	827	25.24	32	3.87
2018	3492	932	26.69	30	3.22
2019	3626	996	27.47	22	2.21
2020	3928	1114	28.36	42	3.78
2021	4511	1445	32.03	41	2.84
2022	4935	1809	36.66	89	4.92
2023	5225	2218	42.45	147	6.63
Total	46503	13979	30.06	595	4.26

Panel F: Summary statistics.

Variable	N	Mean	S.D.	Min.	Median	Max.
Restate	595	0.091	0.288	0.000	0.000	1.000
Same	595	0.346	0.476	0.000	0.000	1.000
Size	595	26.50	2.608	21.58	26.45	31.14
Lev	595	0.683	0.229	0.177	0.706	0.948
Roa	595	0.035	0.043	–0.062	0.017	0.206
Loss	595	0.042	0.201	0.000	0.000	1.000
Ddratio	595	0.387	0.068	0.286	0.364	0.667
Dshsize	595	11.09	3.465	5.000	10.000	21.000
Age	595	2.469	0.644	0.693	2.565	3.434
Big4	595	0.687	0.464	0.000	1.000	1.000
Soe	595	0.571	0.495	0.000	1.000	1.000
AH	595	0.508	0.500	0.000	1.000	1.000
BM	595	0.865	0.229	0.184	0.975	1.249
Tenure	595	6.387	5.093	1.000	5.000	28.000

5.2. Empirical results

5.2.1. Accounting firms simultaneously responsible for ESG report assurance and audit quality

Table 4 presents the regression results regarding whether accounting firms' being responsible for both annual report audit and ESG report assurance enhances audit quality. Column (1) shows the results without controlling for industry and year fixed effects, while Column (2) presents the results with industry and year fixed effects controlled for. In Columns (1) and (2), the coefficients for *Same* are –0.048 and –1.399, respectively, which are significant at the 10 % and 1 % levels. This implies that when ESG report assurance is provided by the same accounting firm responsible for annual report audit, audit quality is enhanced through knowledge spillover effects and reputational effects, thereby supporting the hypothesis of this study. Furthermore, the coefficients for *Big4* are negative and significant at the 5 % level, suggesting that annual reports audited by Big 4 accounting firms are less likely to undergo financial restatements, indicating higher audit quality. The regression results for the other variables are generally in line with expectations.

Table 4
Accounting firms simultaneously responsible for ESG report assurance and audit quality.

Variable	**(1)**	**(2)**
	Restate	***Restate***
Same	–0.048*	–1.339***
	(–1.68)	(–2.60)
Size	0.011	0.501**
	(1.08)	(2.09)
Lev	0.095	5.347
	(1.31)	(1.64)
Roa	–0.408*	–13.608
	(–1.70)	(–1.06)
Loss	–0.045	–1.815
	(–0.93)	(–1.12)
Ddratio	–0.134	–4.580
	(–0.87)	(–0.95)
Dshsize	0.013**	0.044
	(2.55)	(0.52)
Age	–0.063***	–0.554
	(–3.12)	(–1.12)
Big4	–0.055**	–1.721**
	(–2.13)	(–2.47)
Soe	0.094***	1.562***
	(3.60)	(3.22)
AH	0.060**	0.893
	(2.50)	(1.35)
BM	–0.109**	–2.025*
	(–2.18)	(–1.72)
Tenure	0.003	–0.007
	(1.49)	(–0.16)
Constant	–0.134	–12.788***
	(–0.67)	(–2.62)
Observations	595	595
Pseudo R^2	0.12	0.36
Industry FE	NO	YES
Year FE	NO	YES

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

Table 5
Robustness tests.

Panel A: Controlling for different fixed effects and clustering.

Variable	**(1)**	**(2)**	**(3)**	**(4)**
	Controlling for firm fixed effects	**Controlling for accounting firm fixed effects**	**Clustering at the firm level**	**Clustering at the firm and year levels**
	Restate	***Restate***	***Restate***	***Restate***
Same	–0.140*	–2.795***	–1.339**	–1.339***
	(–1.93)	(–3.55)	(–2.01)	(–2.60)
Size	0.146*	1.041**	0.501*	0.501**
	(1.70)	(2.54)	(1.96)	(2.09)
Lev	0.599	10.113**	5.347	5.347
	(1.54)	(2.23)	(1.62)	(1.64)
Roa	–0.831	17.079	–13.608	–13.608
	(–1.21)	(1.33)	(–1.14)	(–1.06)
Loss	0.038	–15.352***	–1.815	–1.815
	(0.39)	(–7.26)	(–1.20)	(–1.12)
Ddratio	–0.318	1.672	–4.580	–4.580
	(–0.99)	(0.29)	(–0.85)	(–0.95)
Dshsize	–0.009	0.026	0.044	0.044
	(–0.90)	(0.18)	(0.39)	(0.52)
Age	–0.378***	–1.579*	–0.554	–0.554
	(–3.71)	(–1.93)	(–0.82)	(–1.12)
Big4	0.156	–27.529	–1.721**	–1.721**
	(1.45)	(–0.00)	(–2.32)	(–2.47)
Soe	0.317***	1.763***	1.562**	1.562***
	(4.82)	(2.70)	(2.56)	(3.22)
AH	–0.004	–0.321	0.893	0.893
	(–0.02)	(–0.30)	(0.99)	(1.35)
BM	0.005	–3.170	–2.025	–2.025*
	(0.03)	(–1.09)	(–1.36)	(–1.72)
Tenure	0.004	0.103	–0.007	–0.007
	(1.04)	(1.54)	(–0.17)	(–0.16)
Constant	–3.269	–2.431	–12.788**	–12.788***
	(–1.56)	(–0.00)	(–2.06)	(–2.62)
Observations	595	595	595	595
Pseudo R^2	0.28	0.55	0.36	0.36
Industry FE	NO	YES	YES	YES
Year FE	YES	YES	YES	YES
Audit FE	NO	YES	NO	NO
Firm FE	YES	NO	NO	NO

Panel B: Controlling for interaction fixed effects.

Variable	**(1)**	**(2)**
	Restate	***Restate***
Same	–0.070*	–0.102***
	(–1.84)	(–2.79)
Size	0.025**	0.039***
	(2.32)	(3.29)
Lev	0.119	0.040
	(1.34)	(0.44)
Roa	–0.112	–0.251
	(–0.36)	(–0.65)
Loss	0.005	0.013
	(0.09)	(0.22)
Ddratio	–0.207	–0.038
	(–0.97)	(–0.17)
Dshsize	0.002	0.009
	(0.39)	(1.53)
Age	–0.037*	–0.047*
	(–1.72)	(–1.85)
Big4	–0.040	–0.073
	(–1.28)	(–1.34)
Soe	0.087***	0.101***
	(2.76)	(3.16)
AH	0.029	0.013
	(0.96)	(0.44)
BM	–0.070	–0.133*
	(–1.15)	(–1.88)
Tenure	0.000	0.005**
	(0.09)	(1.99)
Constant	–0.506**	–0.852***
	(–2.14)	(–3.08)
Observations	595	595
Pseudo R^2	0.33	0.44
Industry FE	YES	YES
Year FE	YES	YES
Industry-Year FE	YES	NO
Audit-Year FE	NO	YES

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

5.2.2. Robustness tests

The impact of the same accounting firm's provision of both annual report audit and ESG report assurance services on audit quality may not stem from the knowledge spillover effects associated with offering both services or an increase in reputational capital. Instead, it could be influenced by characteristics of the accounting firm or individual company. To eliminate these potential confounding effects, we conduct regression analyses while controlling for company fixed effects and accounting firm fixed effects, with the results presented in Columns (1) and (2) of Panel A, Table 5. The coefficients for *Same* are –0.140 and –2.795, respectively, which are significant at the 10 % and 1 % levels, further supporting the conclusions of this study. The data sample spans from 2009 to 2023, consistent with the characteristics of the panel data. Building on the fixed effects controls, and to further mitigate the impact of heteroscedasticity, we employ clustering at the firm and firm-year levels. The results in Columns (3) and (4) indicate that both the coefficients for *Same* are –1.339, significant at the 5 % and 1 % levels, respectively, demonstrating the robustness of the conclusions of this paper.

To further enhance the stability of the results, we additionally incorporate robustness checks by controlling for industry-year interaction fixed effects and accounting firm-year interaction fixed effects. The regression results are presented in Panel B. The findings indicate that even after the inclusion of interaction fixed effects,

Table 6
PSM and Heckman methods.

Variable	**(1)** **PSM**	**(2)**	**(3)** **Heckman**	**(4)**	**(5)**
	Restate	***Restate***	***Same***	***Restate***	***Restate***
Same	–0.668*	–1.475**		–0.875**	–1.642***
	(–1.71)	(–2.53)		(–2.10)	(–3.15)
Size	0.254	0.633**	0.368***	2.361***	3.170***
	(1.21)	(2.10)	(6.58)	(3.94)	(4.02)
Lev	0.932	5.388	–2.886***	–17.145***	–19.005***
	(0.42)	(1.22)	(–4.63)	(–3.04)	(–2.62)
Roa	–19.561**	–19.443	–9.314***	–74.663***	–92.478***
	(–2.31)	(–1.51)	(–2.99)	(–3.59)	(–3.61)
Ddratio	–3.698	–7.059	–1.563	–11.707***	–15.959***
	(–1.12)	(–1.30)	(–1.48)	(–2.93)	(–2.93)
Dshsize	0.111*	0.035	0.056**	0.500***	0.465***
	(1.83)	(0.37)	(2.15)	(3.72)	(3.12)
Age	–0.720**	–0.492	–0.060	–1.067***	–0.967*
	(–2.41)	(–0.82)	(–0.61)	(–3.43)	(–1.96)
Big4	–1.067*	–1.725**	–0.095	–1.503***	–2.181***
	(–1.88)	(–2.18)	(–0.48)	(–2.61)	(–2.85)
Soe	1.340***	1.964***	0.666***	5.117***	6.263***
	(3.48)	(3.56)	(5.03)	(4.36)	(4.37)
AH	0.870*	0.904	0.002	0.896*	0.785
	(1.85)	(1.18)	(0.01)	(1.67)	(1.10)
BM	–2.153**	–1.153	–0.237	–2.491**	–3.320**
	(–1.99)	(–0.60)	(–0.56)	(–2.33)	(–2.34)
Tenure	0.073*	0.027	0.010	0.105**	0.071
	(1.67)	(0.55)	(0.77)	(2.44)	(1.50)
Loss	–1.688	–1.654	–1.073**	–8.939***	–10.615***
	(–1.38)	(–1.07)	(–2.43)	(–3.09)	(–3.61)
IMR				8.800***	10.841***
				(3.63)	(3.79)
Constant	–6.615	–15.872***	–7.994***	–60.617***	–77.697***
	(–1.63)	(–2.72)	(–7.25)	(–4.12)	(–4.19)
Observations	559	559	595	595	595
Pseudo R^2	0.21	0.41	0.24	0.22	0.38
Industry FE	NO	YES	YES	NO	YES
Year FE	NO	YES	YES	NO	YES

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

the provision of both annual report audit and ESG report assurance by the same accounting firm continues to improve audit quality, thereby further supporting the conclusions of this research.

The simultaneous provision of annual report audit and ESG report assurance services by the same accounting firm can enhance audit quality. However, it raises the question of whether firms with higher audit quality are more likely to choose the same accounting firm for both services, which may render the results susceptible to reverse causality. To further mitigate the impact of reverse causality on the regression results, we employ PSM to filter the sample. Using kernel matching, a control group is selected, and the sample undergoes PSM followed by additional testing. Columns (1) and (2) of Table 6 indicate that the coefficients for *Same* are –0.668 and –1.475, respectively, which are significant at the 10 % and 5 % levels. These findings suggest that the results of this study are not merely a product of reciprocal causation.

Furthermore, the decision of a company to choose the same accounting firm for both its annual report audit and ESG report assurance may involve a degree of selection bias. To reduce the potential impact of sample selection bias on the conclusions of this study, we employ the Heckman two-stage estimation procedure. The results of the first stage of the Heckman estimation are presented in Column (3) of Table 6, while Columns (4) and (5) report the results of the second stage. After incorporating the inverse Mills ratio, the coefficient of

Table 7
IV method.

Variable	**(1)**	**(2)**
	Same	***Restate***
Esgbg	0.353*	
	(1.75)	
Same		–17.589*
		(–1.77)
Size	0.084***	1.799**
	(5.18)	(2.10)
Lev	–0.704***	–5.476
	(–4.69)	(–0.78)
Roa	–0.305	–25.846*
	(–0.64)	(–1.92)
Loss	0.073	–0.024
	(0.96)	(–0.02)
Ddratio	0.017	–2.583
	(0.06)	(–0.51)
Dshsize	–0.013	–0.145
	(–1.46)	(–0.94)
Age	–0.007	–0.677
	(–0.25)	(–1.12)
Big4	–0.049	–3.041***
	(–1.03)	(–3.22)
Soe	0.131***	3.469**
	(2.98)	(2.48)
AH	0.015	1.258*
	(0.32)	(1.87)
BM	–0.021	–2.125
	(–0.21)	(–1.49)
Tenure	–0.003	–0.071
	(–1.01)	(–1.25)
Constant	–1.250***	–30.668**
	(–3.16)	(–2.31)
Observations	595	595
Pseudo R^2	0.43	0.35
Industry FE	YES	YES
Year FE	YES	YES

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

Same remains negative and significant, indicating that the results of this study remain robust even after accounting for sample selection bias.

We also employ an IV method to further address the issue of endogeneity. Specifically, the expertise of the accounting firm responsible for the annual report audit in the environmental sector is utilized as an IV for the firm's simultaneous responsibility for ESG report assurance. Drawing on Fan et al. (2013), we measure the accounting firm's expertise in the environmental sector (*Esgbg*) by the ratio of total assets of the accounting firm's clients in heavily polluting industries to the total assets of all clients audited by the accounting firm. An accounting firm that specializes in the environmental industry is more likely to be selected by a company for both its annual report audit and ESG report assurance; however, the impact of the accounting firm's environmental expertise on audit quality is relatively minor, making it suitable as an IV. The results in Column (1) of Table 7 indicate that the coefficient of *Esgbg* is positive and significant, suggesting that accounting firms' possession of environmental expertise indeed encourages companies to select them for ESG report assurance. Additionally, the results in Column (2) show that the coefficient of *Same* remains negative and significant, indicating that the findings of this study remain robust even after the introduction of the IV.

To further mitigate the impact of endogeneity, we first exclude observations that involve the selection of the same accounting firm to provide both ESG report assurance and annual report audit. We then compare the

Table 8
Exclusion of alternative explanations.

Variable	**(1)**	**(2)**
	Restate	***Restate***
Same_diff	0.694	
	(1.10)	
Same_other		–0.092*
		(–1.96)
Size	0.130	0.018
	(0.33)	(1.32)
Lev	7.166	–0.097
	(1.52)	(–0.77)
Roa	–1.212	–0.822***
	(–0.12)	(–2.68)
Loss	–0.279	–0.003
	(–0.20)	(–0.05)
Ddratio	–4.970	–0.348
	(–0.94)	(–1.28)
Dshsize	0.107	0.006
	(0.88)	(0.88)
Age	–0.185	–0.041
	(–0.37)	(–1.65)
Big4	–0.995	–0.049
	(–1.32)	(–1.50)
Soe	1.635**	0.100***
	(2.57)	(2.92)
AH	–0.371	0.096***
	(–0.53)	(2.75)
BM	–0.131	–0.166**
	(–0.09)	(–2.19)
Tenure	0.039	–0.001
	(0.69)	(–0.37)
Constant	–6.861	0.068
	(–0.92)	(0.21)
Observations	389	517
Pseudo R^2	0.39	0.26
Industry FE	YES	YES
Year FE	YES	YES

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

effects of selecting other accounting firms or non-accounting firms for ESG report assurance on audit quality to identify any differences. We define *Same_diff* as follows: if the ESG report assurance is provided by another accounting firm, *Same_diff* takes the value of one; if it is provided by a non-accounting firm, *Same_diff* takes the value of zero. The coefficient of *Same_diff* in Column (1) of Table 8 is not statistically significant, indicating that there is no difference in the impact of selecting other accounting firms or non-accounting firms for ESG report assurance on audit quality. Additionally, to eliminate the concern that the results of this study could be driven by the poorer audit quality associated with companies that select different accounting firms for ESG report assurance, we exclude observations for which ESG report assurance is conducted by other accounting firms. We define *Same_other* as follows: if ESG report assurance is conducted by the same accounting firm as the annual report audit, it takes a value of one; if ESG report assurance is provided by a non-accounting firm, it takes a value of zero. The results from the re-estimation are presented in Column (2) of Table 8. *Same_other* is negative and significant, indicating that the conclusions of this study are not influenced by this potential issue.

To address the small-sample issue, we implement strategies to mitigate the impact of outliers on the conclusions drawn from the analysis. Given the limited sample size, we restrict extreme values of the variables by only regressing data within the 5th to 95th percentiles. This approach reduces the influence of outliers. Addi-

Table 9
Mechanism analysis.

Variable	(1)	(2)
	Delay	*Restate*
Same	5.482***	–1.200**
	(3.74)	(–2.53)
Delay		–0.029*
		(–1.66)
Size	–0.966*	0.442*
	(–1.71)	(1.94)
Lev	6.502	5.372
	(1.15)	(1.54)
Roa	–17.034	–14.668
	(–0.73)	(–0.96)
Loss	–1.588	–1.807
	(–0.50)	(–1.11)
Ddratio	18.421*	–4.220
	(1.79)	(–0.90)
Dshsize	0.496**	0.049
	(2.10)	(0.59)
Age	–3.344***	–0.646
	(–3.36)	(–1.35)
Big4	–4.574**	–1.751***
	(–2.49)	(–2.67)
Soe	3.219***	1.690***
	(2.69)	(3.30)
AH	–8.834***	0.713
	(–5.33)	(1.09)
BM	1.599	–1.531
	(0.35)	(–1.23)
Tenure	–0.190	–0.000
	(–1.44)	(–0.00)
Constant	103.958***	–9.151*
	(7.21)	(–1.92)
Observations	595	595
Adj. R2/Pseudo R^2	0.38	0.37
Industry FE	YES	YES
Year FE	YES	YES

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

tionally, we detect high-leverage points (observations that exert a strong influence on the model) by predicting regression values and examining the residuals and leverage values, subsequently removing these high-leverage points from the analysis. The regression results indicate that the findings remain robust even after accounting for the effects of outliers and high-leverage points. Furthermore, to assess the distribution of coefficients, we randomly exclude 5 % of the samples in each iteration and repeat the main regression 500 times. The results demonstrate that the mean of the coefficient distribution is –1.360, which is closely aligned with the coefficient for *Same* in the baseline regression (–1.339), thus indicating the robustness of our results. Due to space limitations, the tables for the above-mentioned small-sample issue tests are not included, but the relevant results are available upon request.

5.2.3. Mechanism analysis

The previous analysis indicates that the simultaneous provision of annual report audit and ESG report assurance services by the same accounting firm can leverage knowledge spillover effects and reputational capital effects. This dynamic encourages auditors to conduct more thorough verifications and examinations in

Table 10
Impact of internal control quality.

Variable	(1)	(2)
	Restate	*Restate*
Same	–11.486***	–15.862***
	(–3.86)	(–3.83)
Ic	–0.003*	–0.003*
	(–1.85)	(–1.73)
*Same*Ic*	0.015***	0.020***
	(3.97)	(3.71)
Size	0.106	0.247
	(0.53)	(1.04)
Lev	2.721	7.944*
	(1.26)	(1.82)
Roa	–9.093	–3.615
	(–0.95)	(–0.30)
Loss	–1.671*	–1.741
	(–1.66)	(–1.27)
Ddratio	–3.682	–8.480
	(–1.06)	(–1.64)
Dshsize	0.128**	0.160
	(1.98)	(1.41)
Age	–0.384	–0.614
	(–1.29)	(–1.36)
Big4	–0.921*	–1.339*
	(–1.82)	(–1.68)
Soe	1.029***	1.974***
	(2.69)	(3.48)
AH	0.510	0.635
	(1.11)	(0.82)
BM	–1.768*	–1.306
	(–1.82)	(–0.94)
Tenure	0.043	0.024
	(0.96)	(0.46)
Constant	–3.227	–7.823*
	(–0.83)	(–1.83)
Observations	581	581
Pseudo R^2	0.27	0.44
Industry FE	NO	YES
Year FE	NO	YES

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

their practical work, thereby enhancing audit quality through increased audit input. To test the validity of this impact mechanism, we measure audit input using the number of days between the audit report release date and the financial statement date, employing mediation effect analysis. The results are reported in Table 9. Column (1) represents the second step of the mediation effect test, revealing that the coefficient for *Same* is 5.482, which is statistically significant at the 1 % level. This suggests that the concurrent responsibility for the annual report audit and ESG report assurance leads auditors to allocate more resources to audit input. In Column (2), the coefficient of *Delay* is –0.029, significant at the 10 % level, suggesting that the increased audit input acts as a mediator in the relationship between the simultaneous provision of these two services and the improvement of audit quality, demonstrating a partial mediation effect.

6. Additional analysis

6.1. Further tests

6.1.1. Impact of internal control quality

The quality of internal control directly affects the information environment faced by auditors, thereby having a substantial impact on auditing practices. To investigate how the level of internal control influences the

Table 11
Impact of concurrent listing on the Hong Kong Stock Exchange.

Variable	(1)	(2)
	Restate	*Restate*
Same	–2.452*	–3.429**
	(–1.94)	(–2.29)
AH	0.227	0.090
	(0.37)	(0.10)
*Same*AH*	2.315*	2.740*
	(1.83)	(1.72)
Size	0.092	0.341
	(0.51)	(1.43)
Lev	2.554	6.337*
	(1.22)	(1.79)
Roa	–8.966	–10.800
	(–1.04)	(–0.96)
Loss	–1.060	–1.318
	(–0.83)	(–0.93)
Ddratio	–2.719	–5.285
	(–0.88)	(–1.10)
Dshsize	0.159***	0.096
	(2.62)	(1.18)
Age	–0.526*	–0.342
	(–1.74)	(–0.67)
Big4	–0.827	–1.299
	(–1.52)	(–1.49)
Soe	1.280***	1.782***
	(3.52)	(3.51)
BM	–2.033**	–1.945
	(–2.12)	(–1.61)
Tenure	0.042	–0.013
	(0.96)	(–0.29)
Constant	–4.765	–9.906**
	(–1.29)	(–1.98)
Observations	595	595
Pseudo R^2	0.21	0.37
Industry FE	NO	YES
Year FE	NO	YES

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

effect of simultaneously providing two services, we measure the internal control quality of firms using the Dibo Internal Control Index (*Ic*). A higher *Ic* indicates a stronger internal control system. We introduce an interaction term between the level of internal control (*Ic*) and the simultaneous provision of ESG report assurance and annual report audit services (*Same*) in Eq. (1). Higher internal control quality is associated with lower levels of earnings management (Fan et al., 2013; Shi et al., 2024) and reduced information asymmetry faced by auditors. In this context, the knowledge spillover effect between provision of ESG report assurance and annual report audit is diminished. Furthermore, improved internal control enhances auditors' trust in the company (Chalmers et al., 2019), enabling them to directly utilize internal audit data, which reduces their reliance on knowledge related to ESG report assurance and subsequently weakens the enhancement of audit quality. Conversely, when the internal control quality is low, auditors are compelled to invest more reputational

Table 12
Impact of analyst attention.

Variable	**(1)** **Analyst attention**	**(2)**	**(3)** **Analyst report attention**	**(4)**
	Restate	***Restate***	***Restate***	***Restate***
Same	–3.273***	–5.537***	–2.407***	–4.641***
	(–2.77)	(–3.78)	(–2.92)	(–4.26)
Analyst	–0.037**	–0.073***		
	(–2.17)	(–3.04)		
Research			–0.015**	–0.029***
			(–2.49)	(–3.33)
*Same*Analyst*	0.098***	0.140***		
	(3.04)	(3.40)		
*Same*Research*			0.032***	0.051***
			(2.96)	(3.67)
Size	0.288	0.739**	0.300	0.766***
	(1.39)	(2.52)	(1.49)	(2.79)
Lev	2.189	5.989*	2.176	6.248**
	(0.99)	(1.83)	(0.99)	(2.00)
Roa	–9.732	–4.904	–9.331	–5.576
	(–0.96)	(–0.43)	(–0.95)	(–0.48)
Loss	–1.455	–1.058	–1.419	–1.160
	(–1.03)	(–0.68)	(–1.03)	(–0.78)
Ddratio	–2.758	–4.858	–3.215	–5.083
	(–0.88)	(–1.09)	(–0.98)	(–1.10)
Dshsize	0.136**	0.085	0.135**	0.073
	(2.18)	(1.00)	(2.13)	(0.87)
Age	–0.486*	–0.188	–0.584**	–0.214
	(–1.69)	(–0.38)	(–2.00)	(–0.42)
Big4	–1.260**	–2.062***	–1.238**	–1.943***
	(–2.54)	(–2.77)	(–2.40)	(–2.76)
Soe	0.991***	1.498***	1.119***	1.515***
	(2.62)	(2.81)	(2.92)	(2.82)
AH	0.801	0.918	0.794	0.857
	(1.58)	(1.31)	(1.49)	(1.26)
BM	–2.801**	–3.601**	–2.978**	–3.743**
	(–2.32)	(–2.06)	(–2.57)	(–2.32)
Tenure	0.047	0.005	0.041	–0.010
	(1.06)	(0.11)	(0.91)	(–0.21)
Constant	–7.889*	–17.809***	–7.773*	–18.182***
	(–1.78)	(–3.00)	(–1.81)	(–3.24)
Observations	567	567	567	567
Pseudo R^2	0.24	0.40	0.23	0.40
Industry FE	NO	YES	NO	YES
Year FE	NO	YES	NO	YES

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

capital to avoid audit failures. We posit that the positive effect of simultaneously providing annual report audit and ESG report assurance services on audit quality is more pronounced when the level of internal control is lower. Table 10 presents the regression results with the interaction term included, where the coefficients of the interaction terms between *Same* and *Ic* are 0.015 and 0.020, respectively, both statistically significant at the 1 % level. This indicates that the improvement in audit quality from ESG report assurance is more pronounced in samples with lower levels of internal control.

6.1.2. Impact of concurrent listing on the Hong Kong Stock Exchange

The Hong Kong capital market is more mature than that of mainland China, with a higher level of investor protection and a well-established property rights and legal framework. Consequently, the requirements for information disclosure imposed by investors and regulatory authorities are more stringent for companies listed on the Hong Kong Stock Exchange. Firms listed on both the A-share market and the Hong Kong Stock Exchange must adhere to the regulations of the China Securities Regulatory Commission as well as oversight from institutions such as the Hong Kong Stock Exchange (Luo and Wu, 2018), which results in more stan-

Table 13
Economic consequences: audit fee and firm value.

Variable	**(1)**	**(2)**
	Fee	***Tbinq***
Same	0.265***	0.048*
	(3.70)	(1.78)
Size	0.719***	–0.004
	(20.36)	(–0.23)
Lev	–0.641**	0.175
	(–1.99)	(0.86)
Roa	–0.251	3.542***
	(–0.22)	(2.70)
Loss	0.106	0.209*
	(0.79)	(1.65)
Ddratio	–0.164	0.389
	(–0.33)	(1.03)
Dshsize	–0.036***	0.007
	(–2.87)	(0.78)
Age	–0.157***	0.098***
	(–3.85)	(2.70)
Big4	–0.101	0.046
	(–1.18)	(0.68)
Soe	–0.311***	0.063
	(–4.58)	(1.51)
AH	0.536***	–0.058
	(7.68)	(–1.41)
BM	0.043	–2.605***
	(0.22)	(–12.48)
Tenure	0.000	–0.000
	(0.01)	(–0.11)
Constant	–1.649**	2.303***
	(–2.04)	(5.56)
Observations	565	595
Adj. R^2	0.56	0.38
Industry FE	YES	YES
Year FE	YES	YES

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

dardized information disclosure practices. In contrast, mainland China is still in a transitional economic development stage, and compared with the Hong Kong capital market, the judicial system in mainland China is less efficient, and the financial system has not yet matured (Allen et al., 2005). This results in more severe information asymmetry within the A-share capital market. Higher levels of information asymmetry amplify the knowledge spillover effect from providing ESG report assurance. Moreover, due to the more advanced legal system in Hong Kong, auditors must invest more reputational capital to mitigate the risk of audit failure. This inherently high investment in reputational capital diminishes the impact of ESG report assurance on audit quality. To reduce the litigation risks associated with audit failures, auditors engaged with A + H share companies tend to deliver higher-quality audits (Xin and Wang, 2010), which further weakens the knowledge spillover effect of ESG report assurance. Consequently, compared with A + H share companies, the audit quality of firms listed solely on the A-share market is more likely to improve due to the provision of ESG report assurance. Table 11 presents the regression results, where an *AH* value of one indicates that the company is listed on the Hong Kong Stock Exchange, while a value of zero indicates that it is listed only on the A-share market. The coefficients for the interaction terms between *Same* and *AH* are 2.315 and 2.740, respectively, both significant at the 10 % level. This indicates that the enhancement of audit quality from ESG report assurance is more pronounced in samples consisting solely of A-share listed companies.

6.1.3. Impact of analyst attention

Analysts serve as market intermediaries between investors and publicly listed companies by collecting and analyzing information, subsequently disseminating research reports, which alleviates information asymmetry. Companies that attract greater analyst attention typically exhibit higher levels of information transparency and fewer regulatory violations (Li, 2020). This results in lower risk exposure for auditors, thereby diminishing the necessity for them to enhance their reputational capital to avoid audit failures. Furthermore, as the degree of information asymmetry decreases, auditors become less reliant on ancillary information, weakening the knowledge spillover effect associated with ESG report assurance. Consequently, we posit that companies with lower levels of analyst attention experience a more pronounced effect on audit quality when accounting firms provide both annual report audit and ESG report assurance services. We measure analyst attention by the number of analyses conducted by analysts (teams) within a year (with a team counted as one), and measure research report attention by the number of research reports issued within a year. The regression results are presented in Table 12. The outcomes indicate that the coefficients of the interaction terms between *Same* and *Analyst* (*Research*) are 0.098 (0.032) and 0.014 (0.051), respectively, both significant at the 1 % level. This suggests that the enhancement of audit quality through ESG report assurance is more pronounced in firms with lower analyst attention.

6.2. Economic consequences

The previous analysis indicates that when accounting firms simultaneously provide annual report audit and ESG report assurance, they can enhance audit quality through both knowledge spillover effects and reputational capital effects. An increase in audit quality can bolster the bargaining power of accounting firms, allowing them to command higher fee premiums, which manifests as increased audit fees (Palmrose, 1986; Yang et al., 2017). Furthermore, the mechanism analysis in this study reveals that the provision of ESG report assurance contributes to increased audit inputs. As audit inputs rise, so do audit costs, which correspondingly leads to higher audit fees. Based on this premise, we posit that the simultaneous provision of ESG report assurance and annual report audit leads to an increase in audit fees. The results in Column (1) of Table 13 show that the coefficient of *Same* is 0.265, which is statistically significant at the 1 % level. This is evidence that the concurrent provision of ESG report assurance and annual report audit by accounting firms has a positive impact on audit fees.

Previous research demonstrates that high-quality external audits can enhance corporate governance, particularly in the context of China, which is still in a transitional development phase and where legal frameworks remain imperfect. Within this context, external audits can partially substitute for legal mechanisms (Fan and Wong, 2005; Choi and Wong, 2007). Higher audit quality improves the fairness and legitimacy of financial reporting, which can enhance resource allocation efficiency and ultimately lead to value creation for companies (Saito and McIntosh, 2010). In light of this, we posit that the provision of ESG report assurance not only improves audit quality and audit fees, thereby benefiting accounting firms, but also brings advantages to publicly listed companies, as evidenced by an increase in firm value. To test this conjecture, we employ Tobin's Q (*Tbinq*) as a measure of firm value and conduct regression analysis. In Column (2) of Table 13, the coefficient of *Same* is 0.048, which is statistically significant at the 10 % level. This indicates that the simultaneous provision of ESG report assurance can facilitate a win–win situation for accounting firms and publicly listed companies.

7. Conclusion

To investigate whether the provision of both annual report audit and ESG report assurance services by the same accounting firm affects its independence or leads to knowledge spillover effects, and how these may impact audit quality through increased investment in reputational capital, this study empirically examines cases in which the ESG report assurance and annual report audit are provided by the same accounting firm, and finds the following. (1) The provision of annual report audit and ESG report assurance services by the same accounting firm does not compromise the accounting firm's independence; instead, it creates knowledge spillover and reputational effects, thereby enhancing audit quality. (2) Accounting firms that provide ESG report assurance offer enhanced audit quality by increasing audit input. (3) The positive impact of ESG report assurance on audit quality is more pronounced in companies with lower levels of internal control, non-A + H share companies and those with lower analyst attention. (4) In terms of economic consequences, having the same accounting firm responsible for ESG report assurance and annual report audit increases audit fees, while also enhancing firm value.

The findings indicate that the accounting firm responsible for auditing a company's annual report can provide ESG report assurance without compromising its independence; rather, the accounting firm's audit quality is enhanced by knowledge acquisition through the provision of non-audit services and increased reputational capital. Consequently, regulatory authorities do not need to mandate that accounting firms responsible for annual report audit avoid providing ESG report assurance to maintain independence. Furthermore, the economic consequence analysis reveals that the simultaneous provision of ESG report assurance and annual report audit services contributes to an increase in firm value, thus benefiting publicly listed companies. Therefore, it is advisable to encourage listed companies to engage the same accounting firms for both annual report audit and ESG report assurance. This approach not only enhances the assurance of ESG reports and improves the quality of information disclosure but also fosters a win–win situation for accounting firms and publicly listed companies.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Ting Li acknowledges financial support from the National Natural Science Foundation of China (Nos. 72402120, 72432006 and 72372098), the MOE Project of Key Research Institutes of Humanities and Social Science in University (No. 22JJD790093), the 111 Project (No. B18033) and the Fundamental Research Funds for the Central Universities (No. 2022110353). Hongjun Zhu acknowledges financial support from the University Synergy Innovation Program of Anhui Province (No. GXXT-2021-078) and the Humanities and Social Sciences Research Project of Anhui Provincial Department of Education (No. 2022AH010041).

Appendix A. The core content of assurance for ESG reports

(1) *Environmental (E)*

The environmental domain primarily involves aspects such as resource utilization, pollutant emissions, energy consumption and ecological impacts. The key assurance content includes the following:

Category	Assurance content	Key data or indicators
Carbon emissions	Greenhouse gas emissions data (Scope 1, 2 and 3)	Total carbon emissions, carbon emission intensity, carbon neutrality targets
Energy management	Energy consumption and renewable energy utilization	Total energy consumption, proportion of renewable energy
Pollutant emissions	Emission levels of major pollutants	Chemical oxygen demand, nitrogen oxides, sulfur oxides, volatile organic compounds, etc.
Water resource management	Water extraction, usage and discharge	Total water extraction, total discharge, wastewater treatment capacity
Waste management	Management of solid waste and circular economy practices	Amount of municipal and industrial solid waste generated, recycling rates
Environmental compliance	Adherence to regulations and environmental penalties	Environmental compliance incidents, government fine records

(2) *Social (S)*

The social domain focuses on corporate performance regarding labor rights, supply chain management and community relations. The key assurance content includes the following:

Category	Assurance content	Key data or indicators
Employee rights	Labor standards, work environment and safety measures	Working hours, minimum wage, occupational health and safety incident rates
Occupational health	Measures for protecting employee health and safety	Incident rates, work injury rates, health training coverage
Employee development	Education and training, career development, employee satisfaction	Average training hours per employee, promotion rates, turnover rates
Supply chain management	ESG compliance of suppliers	Supplier audit pass rates, proportion of non-compliant suppliers
Customer responsibility	Product quality and consumer rights protection	Product return rates, customer satisfaction scores
Community contribution	CSR activities	Amount of charitable donations, participation in community service events

(3) *Governance (G)*

The governance domain encompasses corporate governance structures, anti-corruption policies and information security. The key assurance content includes the following:

Category	Assurance content	Key data or indicators
Board governance	Company's governance framework and board structure	Proportion of independent directors, establishment of ESG committees
Anti-corruption and compliance	Anti-bribery, anti-money laundering and business ethics	Number of internal whistleblower cases, status of corruption investigations
Information security	Data privacy protection and cybersecurity	Number of data breach incidents, amount invested in cybersecurity
Risk management	ESG risk management mechanisms within the enterprise	ESG risk assessment reports, execution rate of emergency response plans

Case: PwC's assurance of an Energy Company's 2023 ESG report

Key Data:
Environmental (E): Carbon emissions, energy usage, water consumption
Social (S): Accident mortality rate, number of fatalities at work
Governance (G): Number of employees, proportion of female employees

1. ***Assurance Process***

The ESG report assurance was conducted according to the following process:

Step	Description
1. Preliminary communication	Define the scope and standards of assurance—Identify the sources of the company's ESG data.
2. Risk assessment	Identify risks related to data accuracy—Evaluate the existing ESG management system of the company.
3. Data collection	Conduct interviews with management and ESG personnel—Perform on-site inspections of production facilities.
4. Evidence verification	Cross-check greenhouse gas emission data—Review the company's internal environmental compliance reports.
5. Conclusion formation	Compare findings against international standards (ISAE 3000)—Calculate key ESG indicators.
6. Release of assurance report	Provide a limited assurance opinion—Offer recommendations for improvement.

2. ***Degree of Assurance***

Assurance level	Characteristics	Applicable enterprises
Reasonable assurance	In-depth data verification, on-site reviews and interviews; provides a positive assurance conclusion.	Large publicly listed companies with stringent regulatory oversight.
Limited assurance	Relies on self-reported data from the enterprise; primarily based on sampling analyses; offers a negative assurance conclusion.	Smaller enterprises or those disclosing ESG information for the first time.

In this case, PwC applied a limited assurance approach and employed the following assurance methods:

(1) Data verification (validation of the company's greenhouse gas emission reports).
(2) Document review (assessment of the company's ESG management policies).
(3) On-site interviews (discussions with the ESG team and management).
(4) Calculation review (examination of the company's carbon emission calculation methodologies).

3. ***Assurance Findings***

During the review process, PwC identified the following findings:

The company's carbon emission data are accurate; however, there are certain deviations in the calculation methods compared with ISO 14064-3.

Records of occupational health and safety training are comprehensive, yet the training coverage for some frontline employees is relatively low.

The anti-corruption policy is complete; nonetheless, there is a need to enhance the compliance review of third-party suppliers.

4. ***Assurance Report (Sample)***

ESG Assurance Statement
Assurance Provider: Price Waterhouse Coopers (PwC)
Company Name: XXXX
Scope of Assurance: ESG Report for the Year 2023
Assurance Standard: ISAE 3000
Degree of Assurance: Limited Assurance
Assurance Conclusion: Based on our review, we conclude that the ESG report of XX Energy Group is in compliance with applicable disclosure standards in all material respects, and the information presented is accurate, truthful, and free from significant omissions.

References

Allen, F., Qian, J., Qian, M., 2005. Law, finance, and economic growth in China. J. Financ. Econ. 77 (1), 57–116.

Ao, X.B., Sun, L.Y., 2024. Non-audit services and auditor "tone manipulation": empirical evidence based on text analysis. Audit Res. 5, 148–160 (in Chinese).

Beck, P., Frecka, T., Solomon, I., 1998. An empirical analysis of the relationship between MAS involvement and auditor tenure: implications for auditor independence. J. Account. Lit. 7, 65–84.

Beeler, J., Hunton, J.E., 2001. Contingent economic rents: Insidious threats to auditor independence. Working Paper, South Florida University.

Cahan, S.F., Sun, J., 2015. The effect of audit experience on audit fees and audit quality. J. Acc. Audit. Financ. 30 (1), 78–100.

Chalmers, K., Hay, D., Khlif, H., 2019. Internal control in accounting research: a review. J. Account. Lit. 42 (1), 80–103.
Choi, J.H., Wong, T.J., 2007. Auditors' governance functions and legal environments: an international investigation. Contemp. Account. Res. 24 (1), 13–46.
DeFond, M.L., Raghunandan, K., Subramanyam, K.R., 2002. Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions. J. Account. Res. 40 (4), 1247–1274.
Dopuch, N., King, R.R., Schwartz, R., 2001. An experimental investigation of retention and rotation requirements. J. Account. Res. 39, 93–117.
Dou, H., Khoo, E.S., Tan, W., Zhang, J.J., 2024. Superstition, risk aversion, and audit quality: evidence from China. Audit. J. Pract. Theory 43 (4), 51–85.
Fan, J.H., Zhang, Y.M., Liu, Q.L., 2013. Internal control, auditor industry expertise, accruals, and real earnings management. Account. Res. (4), 81–88 + 96 (in Chinese).
Fan, J., Wong, T.J., 2005. Do external auditors perform a corporate governance role in emerging markets? Evidence from East Asia. J. Account. Res. 43, 35–72.
Fang, H.X., Chen, J.J., 2016. Cross-subsidization between two types of audit fees under an integrated model: knowledge spillover effects or economies of scale? Audit Res. (1), 68–75 + 100 (in Chinese).
Frankel, R.M., Johnson, M.F., Nelson, K.K., 2002. The relation between auditors' fees for non-audit services and earnings management. Account. Rev. 77, 71–105.
Gipper, B., Leuz, C., Maffett, M., 2020. Public oversight and reporting credibility: evidence from the PCAOB audit inspection regime. Rev. Financ. Stud. 33 (10), 4532–4579.
Hartzmark, S.M., Sussman, A.B., 2019. Do investors value sustainability? A natural experiment examining ranking and fund flows. J. Financ. 74 (6), 2789–2837.
Hu, X., Hua, R., Liu, Q., Wang, C., 2023. The green fog: environmental rating disagreement and corporate greenwashing. Pac. Basin Financ. J. 78 101952.
Huang, S.Z., 2022. Greenwashing and anti-greenwashing in ESG Reports. Financ. Account. Monthly 1, 3–11 (in Chinese).
Khurana, K., Raman, K.K., 2004. Litigation risk and the financial reporting credibility of Big4 versus non-Big4 audits: evidence from Anglo-American countries. Account. Rev. 79 (2), 473–495.
Krueger, P., Sautner, Z., Tang, D.Y., Zhong, R., 2024. The effects of mandatory ESG disclosure around the world. J. Account. Res. 62 (5), 1795–1847.
Li, K., 2020. Does information asymmetry impede market efficiency? Evidence from analyst coverage. J. Bank. Financ. 118 105856.
Li, Z., Guan, F., Li, Z.Q., 2013. Factors influencing corporate social responsibility report assurance activities: empirical evidence from listed companies in China. Audit Res. 3, 102–112 (in Chinese).
Li, Z., Li, Z.Q., 2012. Do assurance opinions on corporate social responsibility reports have informational value? Empirical evidence from listed companies in China. Audit Res. 1, 78–86 (in Chinese).
Libby, R., Frederick, D.M., 1990. Experience and the ability to explain audit findings. J. Account. Res. 2, 348–367.
Luo, X.X., Wu, L.N., 2018. The impact of capital market openness on corporate auditing: an empirical study based on the "Stock Connect" background. Audit Res. 5, 65–73 (in Chinese).
Maso, L.D., Lobo, G.J., Mazzi, F., Paugam, L., 2020. Implications of the joint provision of CSR assurance and financial audit for auditors' assessment of going-concern risk. Contemp. Account. Res. 37 (2), 1248–1289.
Michas, P.N., Russomanno, D., Zhao, M., 2025. The opportunity for partner industry knowledge sharing within audit offices and audit quality. Rev. Acc. Stud.
Ni, X.Y., Zhang, L.P., 2015. Integrated auditing, audit quality, and audit fees. East China Econ. Manag. 5, 113–117 (in Chinese).
Palmrose, Z.-V., 1986. Audit fees and auditor size: further evidence. J. Account. Res. 24 (1), 97–110.
Pittman, J., Stein, S., Valentine, D., 2023. The importance of audit partners' risk tolerance to audit quality. Contemp. Account. Res. 40 (4), 2512–2546.
Saito, Y., McIntosh, C.S., 2010. The economic value of auditing and its effectiveness in public school operations. Contemp. Account. Res. 27 (2), 639–648.
Shen, H., Lin, H., Han, W., Wu, H., 2023. ESG in China: a review of practice and research, and future research avenues. China J. Account. Res. 16 (4) 100325.
Shi, Y., Zheng, S., Xiao, P., Zhen, H., Wu, T., 2024. ESG performance and cost of debt. China J. Account. Res. 17 (4) 100390.
Simunic, D.A., 1984. Auditing, consulting, and auditor independence. J. Account. Res. 22 (2), 679–702.
Simunic, D.A., Stein, M.T., 1996. The impact of litigation risk on audit pricing: a review of the economics and the evidence. Audit. J. Pract. Theory 15 (2), 119–134.
Stein, S.E., 2019. Auditor industry specialization and accounting estimates: evidence from asset impairments. Audit. J. Pract. Theory 38 (2), 207–234.
Wen, Z.L., Ye, B.J., 2014. Analysis of mediating effects: methodology and model development. Adv. Psychol. Sci. 5, 731–745 (in Chinese).
Xin, Q.Q., Wang, B., 2010. Cross-listing, the big four international firms, and accounting earnings quality. Economic Science 4, 96–110 (in Chinese).
Yang, Q.X., Zhang, J., Yang, Z.Z., Qin, R., 2017. Integrated auditing and audit fees: a dual perspective of scale synergy and knowledge spillover. Account. Res. (3), 82–89+95 (in Chinese).
Yang, S., Liu, Y., Mai, Q., 2018. Is the quality of female auditors really better? Evidence based on the Chinese A-share market. China J. Account. Res. 11 (4), 325–350.
Zhang, J., 2018. Accounting comparability, audit effort, and audit outcomes. Contemp. Account. Res. 35 (1), 245–276.

Zheng, W., Zhu, X.M., Ji, Y., 2015. The level of internal control audit and financial restatement under integrated auditing. Audit Res. 6, 70–77 (in Chinese).

China Journal of Accounting Research 18 (2025) 100436

HOSTED BY

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Using machine learning to identify audit opinion shopping

Jiamei Wang, Chao Yan *

School of Accounting, Zhongnan University of Economics and Law, China

ARTICLE INFO

Article history:
Received 29 May 2024
Accepted 7 July 2025
Available online 28 July 2025

Keywords:
Audit opinion shopping
Machine learning
Auditor change
Abnormal audit fees

ABSTRACT

We select a machine learning model to identify audit opinion shopping and analyze the factors driving the model. To this end, we use six models, namely random forest, gradient boosting decision tree, random undersampling boosting, logistic regression (LR), support vector machine and multilayer perceptron. Among them, LR outperforms the other models. Using game theory, we classify 58 features potentially affecting opinion shopping into audit object, audit subject and audit environment categories. LR is used to obtain each category's importance score. We find that audit object features play a crucial role in audit opinion shopping. We also validate and interpret important features. Finally, we use a model to predict audit collusion. Our paper extends the scope of machine learning to scientifically identify audit collusion risk and reveals important features of audit opinion shopping, which has implications for global audit practice.

1. Introduction

Audit opinion shopping has become a major phenomenon globally and has therefore been widely discussed in both academia and practice. It has been described as a means of collusion between audited firms and auditors to obtain favorable but low-quality audit opinions (Li and Zhao, 2014). Clearly, this behavior seriously harms audit quality. The dissemination of false audit opinions can compromise the transparency of information provided to stakeholders, thereby impeding their comprehension of the financial statements of listed firms. It is obvious that this phenomenon will affect market stability and effectiveness (Wang et al., 2007; Chen et al., 2021; Huang and Duan, 2022), resulting in considerable losses for the capital market and

* Corresponding author.
E-mail address: yanchao420@163.com (C. Yan).

https://doi.org/10.1016/j.cjar.2025.100436
1755-3091/

investors. For example, consider the Enron affair, a major financial fraud case in the United States. Enron's acquisition of audit opinions played a crucial role in its financial fraud, leading to significant losses for numerous investors and causing substantial disruption to the capital market order. In China, the rapid development of the capital market has resulted in a steady increase in the number of listed firms, accompanied by an increase in the financial and business complexity involved. In this context, financial fraud incidents have gradually increased and the number of audit opinion shopping cases has increased rapidly, becoming a focal point of market attention. However, the practice of opinion shopping can compromise audit rigor and adversely impact the professional credibility of the audit discipline (DeFond and Zhang, 2014). Therefore, it is important to issue a scientific warning regarding the purchase of audit opinions. This is of great importance for maintaining audit quality and the credibility of the audit industry. Furthermore, it is essential to protect investor interests and promote market fairness and transparency.

Audit opinion shopping is a collusive behavior between the audited entity and the auditor that is often difficult to identify in practice. The objective of this study is to use a multidimensional model to scientifically identify these collusive phenomena. Implementing a feature importance analysis, followed by further exploration, facilitates the identification of critical factors associated with audit opinion shopping. Our analysis also allows us to investigate the mechanisms through which these factors operate.

Machine learning is an advanced technique that enables computers to learn and make predictions or decisions based on data without being explicitly programmed. Using machine learning models can overcome the strict limitations of linear assumptions in traditional empirical research. In addition, it is easier to intelligently choose functions to adapt to complex and changing data and capture complex relationships between features. Despite the existence of highly nonlinear and challenging-to-intuitively-explain models, accurate out-of-sample predictions can be obtained (Kleinberg et al., 2015). Consequently, this study uses this methodology to construct a multidimensional model to scientifically identify audit opinion shopping and elucidate its underlying mechanisms.

We consider six machine learning models, namely gradient boosting decision tree (GBDT), random forest (RF), random undersampling adaptive boosting (RUSBoost), logistic regression (LR), support vector machine (SVC) and multilayer perceptron (MLP). First, we construct the label variable. According to the literature, there are two methods to purchase audit opinions: changing auditors and paying abnormal audit fees (Du and Guo, 2008; Wu et al., 2013; Kraub et al., 2015; Chen et al., 2016). Therefore, we use two methods to measure audit opinion shopping, based on observations from firms that purchased audit opinions in at least one of two ways. Following this, we develop a framework to summarize and classify the features. Audit opinion shopping can be conceptualized as the dynamic interaction between the audit object and the audit subject concerning issues of audit independence or collusion (Lei, 2004). The nature of this interaction is influenced by the inherent features of both parties involved. Meanwhile, the external environment can also influence both parties, inducing specific behavioral patterns. Studies categorize the factors that influence audit opinion shopping, focusing particularly on the audit object, the audit subject and the audit environment. Accordingly, we construct a comprehensive feature framework, encompassing a total of 58 variables.

After variable construction and feature engineering data processing, the sample is divided into a training set and a test set in a 7:3 ratio. The six models are trained separately on the training set, with parameters being adjusted using 5-fold cross-validation. Subsequently, out-of-sample testing is conducted on the test set. A comparison of the evaluation indicators reveal that the recognition ability of the LR model and the SVC model surpass that of other models. Next, we calculate the relative importance of each feature using the LR model, based on recognition performance and computational speed. In terms of the importance of individual features, six features are identified as particularly important: accounting firm size (*Big4*), current assets (*Currentasset*), operating revenue (*Sale*), total assets (*A*), shareholding ratio of the largest shareholder (*ShrHolder1*) and the degree of real earnings management of the audited firm (*ABEM*). Clearly, the features of the audit object are the most important in terms of identifying audit opinion shopping. Therefore, a baseline model is developed that excludes these features. Next, four subclass features of the audit object are added to explore the incremental contribution of each subclass feature to the model. The results show that in addition to the control variables, each subcategory feature has a substantial incremental effect on the model. This indicates that incorporating all object features significantly enhances the model's recognition ability, suggesting a collaborative impact of these features. For robustness testing, various methodologies are used to divide the dataset, balance the sample and evaluate the importance of different features. Moreover, strongly correlated features are eliminated to verify our main results. In further analyses, we perform an interpretability analysis on the model using SHAP diagrams, thus unveiling the opaque mechanisms underlying the model. We also use feature crossover to analyze the influence of feature combination on the model and the interaction between features. Finally, we develop a model to predict audit opinion shopping behavior.

Our contributions are threefold. First, previous quantitative studies on audit opinion shopping are rather limited, with most studies merely scratching the surface at the descriptive level. In contrast, our study adopts a more sophisticated approach by integrating machine learning models. These models are particularly well suited to the analysis of complex phenomena and the formulation of scientific hypotheses and predictions. As such, we contribute significantly to the body of literature in this particular field.

Second, research in this area predominantly focuses on the impact of a single variable on audit opinion shopping. In contrast, we use a more comprehensive approach, comprehensively analyzing the factors that influence audit opinion shopping. This facilitates an examination of audit opinion shopping behavior from a macro-level perspective. In addition, the ability to conduct quantitative and comparative analyses is essential to determine the relative importance of features from different dimensions. Furthermore, our study goes further by conducting an interpretability analysis. The interpretability of the model is crucial, as it helps to understand the roles played by key factors, shedding light on the underlying mechanisms of audit opinion shopping.

Third, we construct a model for recognizing and predicting audit opinion shopping behavior. The recognition aspect of our model is highly beneficial for regulators, allowing them to quickly identify problematic companies and take timely and appropriate measures to address the issue. In addition, the predictive aspect of our model presents significant value for investors and other stakeholders, allowing them to anticipate potential risks, thereby helping them make more informed decisions and avoid adverse outcomes.

The findings of our study also have important practical implications. First, our study reveals the phenomenon of unfair competition within the audit market. In light of these findings, it is imperative that regulatory authorities enhance their oversight of the audit market. Given the complex nature of audit opinion shopping revealed by our study, regulatory authorities should consider implementing stricter regulations and monitoring mechanisms to ensure fair competition. Second, our model-based approach provides regulatory bodies and the general public with effective tools to identify and prevent different forms of audit opinion shopping. Furthermore, it has the potential to enhance auditor independence by fostering a more transparent and regulated environment, thereby reducing the likelihood of undue influence on audit opinions.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 provides a detailed description of our research design. In Section 4, we present our main results. Section 5 reports the robustness tests. Section 6 provides the results of further analysis and Section 7 concludes the paper.

2. Literature review

2.1. Definition and factors influencing audit opinion shopping

The practice of audit opinion shopping, in which auditees and auditors collude to manipulate audit opinions, is a form of collusive behavior. Firms can obtain favorable audit opinions through specific means, but these audit opinions are of low quality (Li and Zhao, 2014). In the context of audit opinion acquisition, scholars generally focus on audit opinions on financial statements (Wu et al., 2020), particularly on two aspects: how companies purchase audit opinions and the factors that influence this process.

Firms can purchase audit opinions in two main ways. The first method is to change auditors (Zhou and Yao, 2018). The second method is to pay abnormal audit fees (Cao et al., 2013; Li and Zhao, 2014). The decision to change auditors to purchase audit opinions can be divided into two levels: the auditor level and the accounting firm level (Zhang, 2012; Cao and Li, 2019; Chen et al., 2021; Chen et al., 2023). Prior to 2000, scholars generally evaluated audit opinion acquisition by considering the types of audit opinions present before and after the occurrence of an auditor change. In the case of enhanced audit opinions, it was assumed that the company had purchased audit opinions (DeFond and Subramanyam, 1998). Lennox (2000) develops the audit opinion shopping model and measures audit opinion shopping behavior by analyzing the difference in the probability that firms are issued non-standard audit opinions under conditions of auditor change and no change, providing a new method for determining audit opinion shopping. Furthermore, the timing of an auditor's change of position can serve as a signal for audit opinion shopping. A temporary change of auditor will seriously harm the quality of financial statements and cause a negative market reaction (Wang et al., 2024). However, some scholars posit that changing accounting firms is not an effective method to purchase audit opinions (Du and Guo, 2008) and that auditees may also purchase audit opinions by paying abnormal audit fees (Zhang et al., 2019). The purchase of audit opinions, in any manner, is ultimately reflected in audit fees. Simunic (1980) was the first to find that there is a significant correlation between audit fees and audit opinion types. Firms purchase clean audit opinions by increasing audit fees. Subsequently, scholars reveal a positive correlation between abnormal audit fees and improvement in audit opinions (Cao et al., 2013). This finding suggests that firms acquire audit opinions through abnormal audit fees. It is evident that there is a notable audit opinion buying behavior in the Chinese capital market (Jiang and Guo, 2024), which is manifested by either changing auditors or paying abnormal audit fees.

Many factors affect audit opinion shopping. The first factor is supply and demand of purchasing behavior. The game between auditees and accounting firms determines whether audit collusion, that is, the purchase of audit opinions, is necessary. In addition, as a type of market behavior, the free choice of both parties is affected by market features and the regulatory environment. Specifically, firms with high earnings manipulation are motivated to purchase audit opinions to hide their earnings management behavior (Cao et al., 2013; Chen et al., 2021). The hidden motives of enterprises are often driven by either the pursuit of profit or by the need to withstand pressure. Xue et al. (2018) find that when the audited entity has a high concentration of suppliers and customers, firms are more motivated to purchase audit opinions to improve their bargaining power. In the context of the separation of powers, the findings of Quan et al. (2010) suggest a direct correlation between the degree of control exercised by management and the motivation of decision makers to enhance their personal earnings. Equity incentive policies motivate executives to engage in audit opinion shopping (Chen et al., 2015; Chen and Cao, 2018). Li and Zhao (2014) also reveal that enterprises that restated their financial statements in the previous year exhibit a higher propensity to purchase audit opinions. Cao and Li (2019) demonstrate that, following controlling shareholders' equity pledge, the motivation to purchase audit opinions is enhanced, the objective being to reduce the risk of control rights transfer. In addition, financing cost pressure (Qin and Liu, 2019), operating pressure (Xue et al., 2018), delisting pressure (Xie et al., 2010) and follow-up research pressure by analysts and institutional investors (Zhai et al., 2016; Yan et al., 2024) caused by litigation risks may induce the auditee to purchase audit opinions. In addition, Newton et al. (2016) find that the more intense the market competition, the more the auditee is motivated to purchase audit opinions. Of course, there are some factors that can inhibit the purchase of audit opinions. Studying private firms, Li and Zhang (2018) find that the participation of Party organizations in corporate governance can reduce abnormal audit fees and the possibility of audit opinion shopping. Zuo et al. (2013) find that setting up an audit committee and improving its independence can inhibit the purchase of audit opinions. Zhou and Yao (2018) show that media public opinion can monitor firms and reduce their audit opinion shopping behavior. Our study focuses on audit opinion shopping because low-quality audit opinions not only seriously affect the order of the capital market and society but also cause social distrust in auditors and endanger the entire CPA industry (Yuan et al., 2020).

2.2. Applications of machine learning

Machine learning, which relies on a variety of algorithms, has demonstrated its efficacy in constructing complex relationships, efficiently extracting information from high-dimensional data and developing models. These capabilities extend beyond simple measurement of variables and prediction of out-of-sample events, encompassing the areas of causal inference and theory construction (Liu et al., 2023). In financial audits, the integration of machine learning and deep learning has emerged as a dominant trend. Currently, machine learning technology is used in various applications, including variable measurement and event prediction (Athey and Imbens, 2019; Xu et al., 2023). In terms of variable measurement, scholars mainly focus on variables that are difficult to quantify in previous studies, such as executive innovation cognition (Chen et al., 2015), CEO communication style and personality traits (Choudhury et al., 2019; Harrison et al., 2019), manager myopia (Hu et al., 2021), financial statement themes (Brown et al., 2020), corporate culture (Li et al.,

2021) and corporate social responsibility (Fan et al., 2024). Machine learning technology can perform information mining, information extraction and variable construction from unstructured data, even multimodal data, which is a supplement to the existing structured database. For event prediction, current research focuses more on asset price prediction. For example, Li et al. (2017) use a machine learning model to predict the rise and fall of asset prices in the future, while Gu et al. (2020) predict the stock risk premium. However, machine learning methods are gradually being applied to more behavior prediction. For example, corporate fraud (Perols et al., 2017; Xu et al., 2023), financial fraud (Bao et al., 2020) and corporate bankruptcy (Yang and Meng, 2006) have been identified and predicted more accurately by scholars using machine learning models. Furthermore, machine learning can be used in causal inference and theory construction, but there are currently relatively few related studies (Feng et al., 2020; Tidhar and Eisenhardt, 2020), requiring further research. In short, machine learning technology can be widely applied in the field of financial auditing, but current research in this area is insufficient.

3. Research design

3.1. Sample selection

We use China as our research setting for the following reasons. First, China's capital market has developed rapidly in recent years, with a large number of listed firms and diversified businesses and a large number of audit services, which provides a rich sample for the study of audit opinion shopping behavior. Second, China's audit market started late and is still immature. The audit demand is mainly driven by the government, which is a mandatory demand, while the voluntary demand is insufficient. In this market environment, audit opinion shopping behavior has unique manifestations and driving factors and the supply–demand relationship behind it is more diverse and representative, which makes it possible to analyze audit opinion shopping behavior and its driving factors from multiple dimensions. Third, China's legal supervision over audit collusion is not perfect, and the qualitative and quantitative punishment standards for audit opinion shopping are not sufficiently detailed, which may lead to different penalties in actual implementation. This regulatory environment leaves room for audit opinion shopping and provides a unique background for research.

As the disclosure of Chinese institutional investor survey data has been relatively complete since 2014, our sample period begins in 2014. Our feature variables contain information on lagged equity incentives, so our sample period ends in 2021. Finally, we select Chinese A-share listed firms from 2014 to 2021 as our main sample. The institutional investor survey data come from the Wind database, the internal control index is from the DIB database and the data on firms' financial indicators, nature of equity, audit opinions and audit fees come from the CSMAR and CNRDS databases. The following samples are excluded from our core sample: (1) financial firms; (2) samples with missing data and not participating in regression; and (3) firms with only one year of data. To alleviate the effects of extreme outliers, we winsorize the top and bottom 1 % of each continuous variable. Our final sample contains 13,004 observations.

3.2. Variable measurement

The key variable of this paper is audit opinion shopping. As there are two main ways to purchase audit opinions, we measure the purchase of audit opinions from the following two aspects.

3.2.1. Change of accounting firm

Referring to the core idea of the audit opinion shopping model proposed by Lennox (2000), we define firms with the probability of purchasing audit opinions by changing accounting firms and those that actually changed accounting firms as firms that purchased audit opinions. The explanatory variable Q_{it}^{qs} in formula (1) indicates the type of audit opinion obtained by firm i in period t. If it is a non-standard audit opinion, this variable takes a value of 1 and otherwise 0. S_{it} indicates whether a firm has changed accounting firms, taking a value of 1 if so, and otherwise 0. $Controls_{it}$ represent the control variables, including financial leverage, capital profit margin, cash flow from operating activities, internal equity and firm size. Using the regression coefficient of formula (1) by year and by industry, the predicted values Q_{it}^{q1} and Q_{it}^{qo} indicating the audit opinions of firms with or without change of audit firm in the current year can be obtained. Using probit regression, the probability of obtaining non-standard opinions can be obtained. If the probability of "switching audit firms to obtain non-standard opinions" is lower than the probability of "continuing to hire the current firm to obtain non-standard opinions" (that is, $OP = P(Q_{it}^{q1} = 1) - P(Q_{it}^{q0} = 1) < 0$), the firm has an incentive to purchase audit opinions.

$$\hat{Q}_{it}^{qs} = \beta_0 + \beta_1 S_{it} + \beta_2 Controls_{it} + \beta_3 S_{it} Controls_{it} + \beta_4 Q_{it-1} + \beta_5 S_{it} Q_{it-1} + \varepsilon_{it} \tag{1}$$

Previous studies verify the purchase of audit opinions according to formula (2). If the coefficient of OP is negative, it means that listed companies have purchased audit opinions. Therefore, if OP is less than 0 and S_{it} is equal to 1, then the firm has purchased audit opinions by changing accounting firms. Our key variable is *Purchase_change*, which takes a value of 1 if the firm purchased audit opinions and otherwise 0.

$$S_{it} = \lambda_0 + \lambda_1 OP + \lambda_2 Controls_{it} + \varepsilon_{it} \tag{2}$$

3.2.2. Paying abnormal audit fees

For the purchase of audit opinions by paying abnormal audit fees, we refer to Simunic's (1980) audit pricing model and establish a multiple regression model on audit fees, estimating abnormal audit fees using the residual term of the regression model. The specific model for calculating audit fees is provided in formula (3), where Fee_{it} is the audit fees of firm i in year t, while X_{it} is a vector of control variables including firm size, proportion of inventory and accounts receivable, financial leverage, current ratio, return on total assets, growth rate of operating income and other factors affecting audit fees. We take the residual from formula (3) as abnormal audit fees. Specifically, firms whose abnormal audit fees exceed the industry average and whose previous audit opinions were non-standard audit opinions are defined as having purchased audit opinions, *Purchase_abfee*. In this case, *Purchase_abfee* takes a value of 1 and otherwise 0.

$$Fee_{it} = \theta_0 + \sum \theta_i X_{it} + \varepsilon_{it} \tag{3}$$

Overall, our key variable is *Purchase_sum*. If a firm purchases audit opinions by changing accounting firms or paying abnormal audit fees during the current year, the value of *Purchase_sum* is 1 and otherwise 0.

3.2.3. Feature variables

We conduct a comprehensive literature review on the factors that influence audit opinion shopping, finding that scholars have carried out in-depth analyses from diverse perspectives and also involved relevant control variables. Following Lei (2004), we integrate and divide the factors that affect audit opinion shopping into three key dimensions: the audit object dimension, the audit subject dimension and the audit environment dimension.

The features of the audit object can be divided into four categories, namely size, complexity, governance and control variables. First, purchasing audit opinions is a game between the auditee and the auditor. The power asymmetry between the two parties puts pressure on the auditor, leading to a loss of independence. Therefore, the size of the auditee may affect audit collusion (Fang et al., 2020). Due to the diversity of indicators for assessing firm size in different industries and in different life cycles, we use five indicators to measure the size of the audited entity, namely total assets, operating revenue, profits, net assets and current assets. Second, from the perspective of motivation, under China's highly centralized ownership structure, if there are defects in corporate governance, it will easily lead to financial fraud and the purchase of audit opinions (Li and Ren, 2012). The governance structure of listed companies is an important institutional arrangement to resolve information asymmetry and protect the legitimate rights and interests of investors. We use 15 indicators to measure governance structure, including the proportion of independent directors, executive power, Party organizations' participation in corporate governance and the nature of property rights. The complexity of audit activities will continue to increase, increasing the difficulty of auditing. To some extent, the complexity of the auditee will also increase the motivation for fraud and induce firms to engage in audit collusion. For example, firms' status in the supply chain may affect their level of motivation to purchase audit opinions

(Xue et al., 2018). We use eight indicators to measure the complexity of the audit object, including the proportion of accounts receivable and inventory, the concentration of suppliers and customers, the degree of earnings management and financial statement restatement. In addition, we control for five variables, including whether the firm is loss-making, the province where the office is located, the year of establishment of the firm and internal control. Among them, better internal control can promote external supervision, thereby inhibiting the purchase of audit opinions by firms (Wang et al., 2021).

Regarding the features of audit subjects, we divide them from the perspective of cost. For firms, rent-seeking demands often come with excess returns, but whether they are accepted depends on the costs incurred by firms. The term "cost" here refers to a broad concept that includes both tangible and intangible costs. We mainly consider reputation, risk and new customer acquisition, which will affect a firm's decision to engage in audit opinion shopping behavior. Therefore, we divide the features of audit subjects into four categories, namely firm size, audit risk premium, firm expertise and audit results. Accounting firm size is usually positively correlated with audit quality (Wu et al., 2015; Xu and Qi, 2023). The larger the accounting firm, the more attention it usually pays to its reputation and market position. As such, it tends to comply with industry norms and professional ethics and to avoid participating in audit collusion and other violations. In addition, the accounting firm will assess the risk of a potential client firm before agreeing to audit it and will choose whether or not to audit the firm according to the assessment results. Therefore, the risk taken by the firm will also affect audit opinion shopping behavior. We use the risk of six audited units to measure the risk taken by the auditor, representing the cost of the risk faced by the auditor. The ability to acquire new customers can be reflected in a firm's market share. Firms with a higher market share can find new customers at lower cost and thus benefit from greater independence. Auditors with strong industry expertise can easily find new customers in the same industry, which improves their independence to some extent, thereby inhibiting the purchase of audit opinions (Ji and Zhang, 2019). We measure audit expertise by industry and region. At the same time, as the ultimate goal of audit opinion shopping, audit results are inseparable from this behavior, so we also classify the types of audit opinions into the features of audit subjects.

Regarding the features of the audit environment, Lei (2004) argues that financial fraud and audit collusion coexist, with their symbiotic environment including the market environment, social environment and cultural environment. Because the cultural environment and the social environment are closely related and difficult to define clearly, we combine them under the social environment and divide the social environment into the legal environment and the monitoring environment. We use 15 indicators to measure the audit environment, including the audit market, the legal environment and the monitoring environment. The audit market includes the degree of marketization and the degree of market competition of firms. As a type of market behavior, the purchase of audit opinions will be affected by the degree of marketization. At the same time, the pressure brought by market competition will affect firms' decision-making. The legal environment is measured by the legal environment index, which has an impact on illegal acts. The monitoring environment includes the media, investors and analysts. A large number of studies have shown that external supervision has an impact on the purchase of audit opinions (Zhai et al., 2016; Zhou and Yao, 2018; Yan et al., 2024). The specific classification of each feature category is shown in Table 1 and definitions are provided in Appendix A.

3.3. Sample distribution and descriptive statistics

The sample distribution is shown in Table 2. From 2014 to 2021, on average 5.61 % of the observations in our sample purchased audit opinions each year. Compared with paying abnormal audit fees, more firms choose to purchase audit opinions by changing accounting firms. However, since 2018, the number of firms purchasing audit opinions by paying abnormal audit fees has shown an upward trend, which may be due to the fact that it is easier to find ways to change accounting firms, so firms are gradually turning to more covert payment behavior.

Table 3 shows the descriptive statistics. The results show that the average value of *Purchase_sum* is 0.056, indicating that 5.6 % of the observations in our sample purchased audit opinions. From the perspective of target value, the sample is unbalanced. The average value of *Purchase_change* is 0.043, while the average value of *Purchase_abfee* is 0.018 and the sum of the two is greater than 0.056. The reason is that some samples not only changed accounting firms but also paid abnormal audit fees. Our main target variable *Purchase_sum* is defined

Table 1
Features variables.

Category	Subcategories	Feature name	Variable
Audit object	Size	Total assets	*A*
		Sales revenue	*Sale*
		Net interest	*Netinterest*
		Net assets	*NetA*
		Current asset	*Currentasset*
	Complexity	Accounts receivable ratio	*AR_A*
		Inventory ratio	*INV_A*
		Supply chain relationships	*T5cusoirt*
		Customer relationships	*T5suplpart*
		Accrued earnings management	*Absem_DACC*
		Real earnings management	*ABEM_Proxy*
		Restatement of financial statements	*Restate*
			Restate0
	Governance structure	Ratio of independent directors	*InDrcRat*
		Shareholding ratio of the largest shareholder	*ShrHolder1*
		Controlling shareholders' equity pledge	*Pld_dum*
		Participation of Party organizations in corporate governance	*Party_per*
			Party_per_IsMTMT
			Party_per_IsMTB
			Party_per_IsSupervisor
			Party_CG
		Executive power	*GmShrRat*
			Manage_female
			Hubris
		Nature of property rights	*SOE*
		Equity incentive	*ESO*
			ESO0
		Board of auditors	*IsAuditCommitteeSetUp*
	Control variable	Loss or not	*Loss*
		Province of office	*Province*
		Firm age	*Firm_age*
		Internal control	*IC*
			Deficiency_Type
Audit subject	Size	Big four	*Big4*
	Risk premium	Operational risk of the audited entity	*Beta*
		Delisting risk of audited entity	*Stpt*
		Debt service risk of the audited entity	*LongLia_A*
		Financial risk of the audited entity	*Financial_Risk*
		Liquidity risk of the audited entity	*Current_ratio*
		Litigation risk of the audited entity	*Litigation_Risk*
	Specialty	Area expertise	*AreaR*
		Industry expertise	*MSR*
	Audit result	Audit opinion on the annual report	*Audittyp_score*
Audit environment	Audit market	Marketization level	*Market*
		Market competition	*HHI*
	Legal environment	Legal environment	*Law_index*
	Monitoring environment	Newspaper media coverage	*J_F_News*
		Online media coverage	*J_F_Online*
		Investor attention	*Investor_neg*
			Investor_pos
			Sum_Readnum
			Sum_Commentnum
		Analyst tracking	*AnaAttention*
		Research report tracking	*ReportAttention*
		Institutional investor	*InsInvestorProp*
			Institutenumber
			Institutenumber_sum
			Is_Institute

as the purchase of audit opinions in one or both ways. The descriptive statistics of the other feature variables are basically the same as those of previous studies.

3.4. Machine learning model

We mainly use three integrated learning models, namely GBDT, RF and RUSBoost, and three traditional machine learning models, namely LR, SVC and MLP, to identify audit opinion shopping behavior. To fully

Table 2
Sample distribution.

Year	*Purchase_sum*	*Purchase_change*	*Purchase_fee*	Observation	Proportion
2014	100	76	30	1806	5.54 %
2015	72	58	19	1568	4.59 %
2016	121	100	29	1952	6.20 %
2017	59	48	15	1332	4.43 %
2018	74	64	15	1487	4.98 %
2019	102	75	33	1501	6.80 %
2020	97	67	35	1442	6.73 %
2021	108	65	53	1916	5.64 %
Total	733	553	229	13,004	5.61 %

Table 3
Descriptive statistics.

Variable	Observation	Mean	SD	Min	Max
Purchase_change	13,004	0.043	0.202	0.000	1.000
Purchase_abfee	13,004	0.018	0.132	0.000	1.000
Purchase_sum	13,004	0.056	0.231	0.000	1.000
A	13,004	22.561	1.281	19.997	26.382
Sale	13,004	21.849	1.448	18.615	25.859
Netinterest_ratio	13,004	0.046	0.175	−0.981	0.451
NetA	13,004	21.826	1.174	18.287	25.219
Currentasset	13,004	21.863	1.281	19.196	25.680
AR_A	13,004	0.126	0.103	0.000	0.473
INV_A	13,004	0.144	0.134	0.000	0.717
T5cusoirt	13,004	0.298	0.211	0.011	0.960
T5suplpart	13,004	0.330	0.195	0.045	0.923
Absem_DACC	13,004	0.058	0.062	0.001	0.368
ABEM_Proxy	13,004	0.009	0.211	−0.794	0.570
Restate0	13,004	0.202	0.402	0.000	1.000
Restate	13,004	0.199	0.399	0.000	1.000
InDrcRat	13,004	0.387	0.101	0.000	0.667
ShrHolder1	13,004	0.328	0.144	0.090	0.730
Pld_dum	13,004	0.472	0.499	0.000	1.000
Party_per	13,004	0.118	0.157	0.000	0.652
Party_per_IsMTMT	13,004	0.121	0.222	0.000	1.000
Party_per_IsMTB	13,004	0.142	0.184	0.000	0.667
Party_per_IsSupervisor	13,004	0.137	0.228	0.000	1.000
Party_CG	13,004	0.603	0.489	0.000	1.000
GmShrRat	13,004	0.055	0.118	0.000	0.559
SOE	13,004	0.366	0.482	0.000	1.000
ESO	13,004	0.105	0.307	0.000	1.000
ESO0	13,004	0.108	0.310	0.000	1.000
IsAuditCommitteeSetUp	13,004	0.991	0.096	0.000	1.000
Loss	13,004	0.132	0.338	0.000	1.000
Firm_age	13,004	3.142	0.229	1.946	4.159
Financial_Risk	13,004	2.210	0.876	1.000	3.000
Litigation_Risk	13,004	0.278	0.448	0.000	1.000
Deficiency_Type	13,004	0.959	1.376	0.000	3.000
IC	13,004	6.158	1.623	0.000	8.141
Big4	13,004	0.057	0.232	0.000	1.000
Beta	13,004	1.102	0.335	0.271	1.938
Stpt	13,004	0.020	0.139	0.000	1.000
LongLia_A	13,004	0.080	0.092	0.000	0.414
Current_ratio	13,004	1.857	1.465	0.297	10.406

(continued on next page)

Table 3 (*continued*)

Variable	Observation	Mean	SD	Min	Max
Market	13,004	9.713	1.613	4.448	12.390
HHI_B	13,004	0.144	0.223	0.013	1.000
Audittyp_score	13,004	1.063	0.332	1.000	4.000
Law_index	13,004	11.099	3.112	2.651	16.507
J_F_News	13,004	0.386	0.413	−1.000	1.000
J_F_Online	13,004	0.145	0.219	−0.511	0.636
Investor_neg	13,004	0.217	0.043	0.061	0.394
Investor_pos	13,004	0.286	0.050	0.134	0.599
Sum_Readnum	13,004	15.942	0.998	13.713	18.095
Sum_Commentnum	13,004	9.223	1.044	6.759	11.882
AnaAttention	13,004	1.422	1.176	0.000	3.784
ReportAttention	13,004	1.775	1.482	0.000	4.727
InsInvestorProp	13,004	0.447	0.237	0.005	0.911
Is_Institute	13,004	0.495	0.500	0.000	1.000
Institutenumber	13,004	1.410	1.765	0.000	5.663
Institutenumber_sum	13,004	0.788	0.955	0.000	3.401
Hubris	13,004	0.629	0.183	0.286	1.000
Manage_female	13,004	0.162	0.163	0.000	0.667
AreaR	13,004	0.122	0.166	0.000	0.641
MSR	13,004	0.061	0.056	0.001	0.264

use our sample data and in line with the actual situation of the capital market, we refer to Chawla et al. (2002) and Wu and Chen (2023) and use synthetic minority over sampling technology (SMOTE) to solve the sample imbalance problem. SMOTE is a method that uses oversampling technology to deal with imbalanced datasets. It increases the number of minority classes by synthesizing new minority samples, to make the dataset balanced. At the same time, thanks to oversampling technology, we can reuse fraud samples multiple times. This process allows the model to repeatedly use and thoroughly study relevant feature indicators, thus enhancing the influence of these feature indicators in model construction. It also preserves as much detail as possible about the sample features, so that the model can better understand the correlation between audit opinion shopping samples and non-audit opinion shopping samples and features. Referring to the literature (Zhou et al., 2022; Zhang et al., 2023), we divide the training set and the test set at a ratio of 7:3. First, the training set is used to train the samples, and then 5-fold cross-validation is used to adjust the parameters. Finally, the test set is used for testing, with the target value of the test set being the output and the target value being compared with the actual value. The model with the best recognition effect is identified by calculating and comparing the evaluation indices. To ensure consistency in our training results, we set the *random_state* parameter to 0 and standardize the data.

4. Empirical results

4.1. Main results

When evaluating the performance of classification models, five key indicators are commonly used, namely Accuracy, Precision, Recall, F1 value and AUC. Accuracy is the proportion of correctly identified samples. Because our sample is unbalanced, if the model simply identifies all samples as firms that did not purchase audit opinions, its accuracy can reach 94.4 %. This clearly shows that in the case of unbalanced sample distribution, the Accuracy index cannot fully and accurately reflect the real performance of the model, so it cannot be used as a reliable basis for evaluating the model. The Precision rate measures the proportion of correctly identified samples that are identified as target values, while the Recall rate reflects the proportion of correctly identified individuals that are actually target samples. The F1 value is the harmonic average of Precision and Recall, which should be considered holistically. However, due to the small proportion of audit opinion shopping observations in the total sample, the model will identify a large number of non-purchase samples as purchase samples while identifying samples that actually purchase audit opinions, leading to a low Precision value. When the difference between Precision and Recall is large, the harmonic average pays more attention to the influence of smaller values, so the evaluation effect of the F1 value will be greatly reduced. As the F1 value may be limited by the classification threshold and sample distribution, we follow previous studies (Bao et al., 2020; Bertomeu et al., 2021; Wu and Chen, 2023) and select AUC as the model evaluation index. AUC refers to the area surrounded by the coordinate axis under the ROC curve. The ROC curve shows the classification ability of the binary classification model. The closer the curve is to the upper left corner, the better the classification effect of the model. The AUC value range is between 0 and 1. The closer the AUC value is to 0 or 1, the better the classification effect of the model. When the AUC value is equal to 0.5, it means that the classification ability of the model is the same as that of random guessing. At this point, the index loses its practical application value in model evaluation. In summary, we select Precision, Recall and AUC as the three evaluation indicators to comprehensively and objectively evaluate the recognition effect of the model.

In addition, we refer to Bao et al. (2020) and use Precision@k, Recall@k and NDCG@k to evaluate the recognition efficiency of the model. Precision@k refers to the proportion of correct relevant results identified among the top k results with the highest probability of purchasing audit opinions produced by the model. This indicator measures the Precision rate of the model among the top k results most likely to purchase audit opinions. Recall@k refers to the proportion of results identified as audit opinion shopping among the first k results of all samples that actually purchased audit opinions. This indicator measures the recall rate of the model in the first k results. NDCG@k (normalized discounted cumulative gain) is an evaluation index considering the return order, which is used to evaluate the ranking quality of the recommendation system or search engine when returning the top k results. NDCG@k considers both the relevance and the order of the results. In other words, the most relevant results should be ranked higher. If the model can accurately rank the results more likely to purchase audit opinions at the top, its ranking quality is higher, with higher NDCG@k values indicating that the model performs better at sorting. To calculate the efficiency index, we first need to determine the value of k. For further information, refer to Bao et al. (2020). We set the value of k based on the number of target samples in the test set. The proportion of target samples in the test set is about 5.90 % and the number of samples in the test set is 3901. Therefore, when calculating the efficiency index, k is rounded to 230. Similarly, when distinguishing the purchase method of audit opinions, based on the proportion of target samples, the value of k is set to 170 and 80, respectively.

The AUC value is a relatively comprehensive index, which is not affected by the model threshold and is suitable for evaluating the classification effect of imbalanced samples. Therefore, we select AUC as the primary evaluation index for model efficiency comparison (Wu and Chen, 2023). In addition, our goal is to identify the purchase of audit opinions, which constitutes negative behavior, the omission of which can lead to serious consequences. Therefore, it is more important to identify samples that actually purchase audit opinions than firms that do not. Therefore, we pay more attention to the value of Recall than that of Precision. Because the proportion of audit opinion shopping samples in the total sample is small, the model will identify samples that actually purchase audit opinions and a large number of non-purchase samples as purchase samples, which will lead to a low Precision value. In this paper, the sample is divided into a training set and a test set at a ratio of 7:3[1] and we use SMOTE[2] sampling to deal with the imbalanced data. The model evaluation indices are calculated by comparing the model output results on the test set with the actual results, and the results are listed in Table 4. The models listed in Panel A of Table 4 concern the purchase of audit opinions in two ways, namely changing firms and paying abnormal audit fees. Regarding the AUC index, the scores for LR and SVC are 0.7938 and 0.7914, respectively, indicating that the classification effect of these models is significantly

[1] To enhance the robustness of our results, we change the division of our training and test sets. First, we divide the dataset by year, designating the 2014–2019 period as the training set and the 2020–2021 period as the test set. Second, we divide the full sample at a ratio of 8:2. The comparison results show that the LR and SVC models still perform well. The specific results are presented in Appendix B.

[2] Our dataset is imbalanced. To solve the problem of imbalanced sample categories, we use the SMOTE oversampling method. However, we also test other methods, including random undersampling, mixed sampling and weighting of a few categories. A comprehensive comparison shows that the LR model still performs relatively well. The specific results are presented in Appendix C.

Table 4
Model evaluation.

Model	Precision	Recall	AUC	Precision@k	Recall@k	NDCG@k
Panel A: *Purchase_sum*						
RF	0.4615	0.2328	0.6078	0.3826	0.3793	0.4020
GBDT	0.4267	0.2759	0.6262	0.3652	0.3621	0.3831
RUSB	0.3279	0.3448	0.6501	0.3348	0.3319	0.3416
LR	0.2096	0.7716	0.7938	0.3913	0.3879	0.4123
SVC	0.1974	0.7845	0.7914	0.3913	0.3879	0.4147
MLP	0.3843	0.4009	0.6801	0.3826	0.3793	0.4191
Panel B: *Purchase_change*						
RF	0.3974	0.1813	0.5843	0.3000	0.2982	0.3229
GBDT	0.3333	0.2281	0.6036	0.3000	0.2982	0.3532
RUSB	0.2537	0.2982	0.6290	0.2765	0.2749	0.2773
LR	0.1723	0.7719	0.8010	0.3294	0.3275	0.3415
SVC	0.1681	0.7953	0.8075	0.3176	0.3158	0.3297
MLP	0.2882	0.2865	0.6271	0.2882	0.2865	0.3326
Panel C: *Purchase_fee*						
RF	0.4490	0.2651	0.6290	0.3875	0.3735	0.4046
GBDT	0.4154	0.3253	0.6577	0.3750	0.3614	0.3583
RUSB	0.2824	0.2892	0.6366	0.3000	0.2892	0.3617
LR	0.1440	0.6747	0.7938	0.4750	0.4578	0.4887
SVC	0.1351	0.6627	0.7852	0.4375	0.4217	0.4771
MLP	0.4521	0.3976	0.6936	0.4125	0.3976	0.4634

Note: (1) Parameters for each model: For the RF model, the values of max_features, min_samples_leaf, max_depth, min_samples_split, and n_estimators are set to 8, 2, 40, 2, and 340, respectively. For the GBDT model, the learning_rate and n_estimators are set to 0.1 and 260, respectively. For the RUSBoost model, the learning_rate and n_estimators are set to 1 and 260, respectively. For the LR model, the solver, C, and penalty are set to "liblinear", 0.1, and "l1", respectively. For the SVM model, the C, gamma, and kernel are set to 100, 0.1, and "linear", respectively. For the MLP model, the hidden_layer_sizes, learning_rate, max_iter, activation, alpha, learning_rate_init, solver, and batch_size are set to (290,), "constant", 100, "tanh", 0.0001, 0.001, "adam", and 70, respectively. (2) In Panel A, the value of k is 230; in Panel B, k is 170; and in Panel C, k is 80. The selection of the k value is primarily based on the proportion of target samples in the test set.

higher than that of other models. Observing the Recall rate index, we find that the Recall rates of the LR and SVC models are 0.7716 and 0.7845, respectively, indicating that 77.16 % and 78.45 % of the auditees in the audit opinion shopping sample are correctly identified. Combined with the confusion matrix of the LR model test results, we find that among 232 target samples in the test set, 179 samples that purchased audit opinions are accurately identified. However, in terms of Precision, the performance of these two models is no better than that of the other four models, or even worse. In terms of model efficiency, Precision@k and Recall@k for LR and SVC are higher than those for the other models, but NDCG@k is slightly lower than that of the MLP model. Panel B and Panel C of Table 4 distinguish between methods of purchasing audit opinions. Panel B models audit opinion shopping behavior by changing audit firms, while Panel C models it paying abnormal audit fees. Finally, the results of these tests on the test set are listed in Table 4, which are basically consistent with those in Panel A. In summary, we suggest that the LR and SVC models are better than the other four models in recognizing audit opinion shopping behavior.

4.2. Feature importance

Considering the AUC and Recall values, overall efficiency index, calculation cost and speed, we suggest that the LR model has better comprehensive performance in identifying audit opinion shopping behavior. In terms of training speed, the LR model is generally faster than the SVC model, because the computational

Table 5
Feature importance.

Category	Feature	Feature importance
Panel A: Individual feature importance		
Audit subject	*Big4*	15.21 %
Audit object	*Currentasset*	14.33 %
Audit object	*Sale*	10.88 %
Audit object	*A*	8.83 %
Audit object	*ShrHolder1*	8.36 %
Audit object	*ABEM_Proxy*	8.24 %
Audit object	*Absem_DACC*	3.78 %
Audit object	*NetA*	3.63 %
Audit subject	*Audittyp_score*	3.17 %
Audit environment	*InsInvestorProp*	2.67 %
Panel B: Sum of feature importance of each category		
Audit object		67.52 %
Audit subject		23.25 %
Audit environment		9.23 %
Panel C: Sum of feature importance of each subcategory		
Audit object	Size	37.78 %
	Complexity	15.21 %
	Governance structure	13.49 %
	Control variable	1.03 %
Audit subject	Size	15.21 %
	Risk premium	4.49 %
	Audit result	3.17 %
	Specialty	0.38 %
Audit environment	Monitoring environment	7.48 %
	Audit market	1.65 %
	Legal environment	0.11 %

complexity of the LR model is low and its training process is relatively simple. Therefore, with an equivalent recognition effect, we choose the more efficient LR model. Based on the LR model, we obtain the importance ranking of each feature and the cumulative importance ranking of each feature.[3] The results are presented in Table 5, with Panel A listing mainly the top 10 feature variables in order of importance. According to the results, accounting firm size (*Big4*), current assets (*Currentasset*), operating revenue (*Sales*), total assets (*A*), the shareholding ratio of the largest shareholder (*ShrHolder1*) and the degree of real earnings management (*ABEM_Proxy*) of the audited entity are the top six features, with an importance score of more than 8 %, indicating that these features are important in identifying the audit opinion shopping behavior of firms.

Furthermore, we summarize the importance of the features in Panel B of Table 5. The results show that the audit object features are the most important in identifying audit opinion shopping behavior, while the importance of the audit subject and audit environment features is less. In Panel C of Table 5, we analyze the importance of each subcategory feature and find that among the audit object features, the audited object size feature is the most important, followed by the complexity and governance structure features, while the control variable feature is less important. At the same time, the importance of audit firm size reaches 15.21 % in the audit subject features, which shows that the size of audit subjects is very important in identifying audit opinion

[3] As the LR model does not have a feature importance parameter, we cannot directly obtain the importance score of each feature. Thus, we use the "coef_" parameter of the model to obtain the weight of each feature and measure its impact on the recognition results.

Table 6
Incremental effect of features.

Model	Precision	Recall	AUC	Precision@k	Recall@k	NDCG@k
Base	0.1398	0.7241	0.7212	0.3217	0.3190	0.3490
Base + Size	0.1458	0.7026	0.7212	0.3304	0.3276	0.3636
Base + Complexity	0.1842	0.8017	0.7886	0.3739	0.3707	0.3864
Base + Governance	0.1589	0.7026	0.7337	0.3217	0.3190	0.3361
Base + Control	0.1410	0.7284	0.7239	0.3217	0.3190	0.3532
Base + Size + Complexity	0.1874	0.7845	0.7847	0.3870	0.3836	0.4118
Base + Size + Governance	0.1686	0.6983	0.7403	0.3565	0.3534	0.3817
Base + Size + Control	0.1517	0.7284	0.7355	0.3304	0.3276	0.3623
Base + Complexity + Governance	0.2011	0.7586	0.7841	0.3739	0.3707	0.3885
Base + Complexity + Control	0.1850	0.7888	0.7846	0.3870	0.3836	0.3989
Base + Governance + Control	0.1649	0.7198	0.7447	0.3304	0.3276	0.3499
Base + Size + Complexity + Governance	0.2082	0.7629	0.7898	0.3870	0.3836	0.4078
Base + Size + Complexity + Control	0.1881	0.7931	0.7884	0.3696	0.3664	0.3975
Base + Size + Governance + Control	0.1672	0.7026	0.7407	0.3565	0.3534	0.3735
Base + Complexity + Governance + Control	0.2044	0.7586	0.7860	0.3870	0.3836	0.3998
Base + All	0.2096	0.7716	0.7938	0.3913	0.3879	0.4123

Note: The target variable of the model is *Purchase_sum*. The dataset is split into training and test sets in a ratio of 7:3, and thus the value of k is taken as 230.

shopping behavior. Among the audit environment features, the importance of the supervisory environment features is also relatively high, at 7.48 %, which verifies the importance of previous research on analysts, institutional investors and media reports.

4.3. Incremental effect of features

Using the LR model, we examine the incremental effect of each subclass feature on the classification performance of the model by gradually adding each subclass feature of the audit object. First, the baseline model is established, which only includes two categories of features: the audit subject and the audit environment. Next, we add the features of audit object subcategories, including size, complexity, governance structure, control variables and permutation and combination of these four types of features based on the baseline model to create 15 "Base +" models. The above 16 models are retrained and the test set is identified. The results are shown in Table 6. We find that adding subcategory object features to the baseline model improves the AUC index of other models compared with the baseline model, except when adding size features. Especially when all object features are added, the AUC score of the model on the test set is the highest, at 0.7938, indicating that there is a joint effect among the subcategory features. When the size and governance features are added to the baseline model, the Recall rate is lower than that of the baseline model, which indicates that the recognition ability of the model for real samples of audit opinion purchases is weaker. However, from the perspective of Precision rate, adding these two types of features improves the model compared with the baseline model, which indicates that the model also reduces the misidentification of non-purchase samples. Combined with the AUC value, adding the governance features increases the AUC value of the baseline model from 0.7212 to 0.7337, indicating that these features improve the recognition effect of the model. However, the AUC value of the size features does not change compared with the baseline model, but this does not mean that they are not important. Indeed, we can see that most of the evaluation indices improve after adding the size features to the "Base + Governance" model. After adding the size features to the "Base + Complexity + Governance + Control" model, each evaluation index also improves. This shows that the incremental contribution of the size features to the model is generated jointly with other features.

5. Robustness tests

5.1. Significance of the incremental effect

In the empirical results, we find that certain audit object features have incremental effects on the identification of audit opinion shopping behavior. Furthermore, we explore the significance of the incremental effects of adding four subcategory features to the baseline model. Therefore, referring to Xu et al. (2023), we retrain and retest the model in different random states by setting *random_state* to 0–29. Finally, we obtain the test results of 30 model groups in different random states and test the significance of the incremental effect through the mean difference between groups.

The results are shown in Table 7. The results for Precision and AUC are roughly the same. By adding three categories of features, namely size, complexity and governance, to the baseline model, the incremental effect of the model is highly significant and the p-values are all less than 0.01. The incremental contribution of the control variable features to the model is relatively small and not significant. Looking at the Recall rate in Panel B, we see that only the complexity features contribute incrementally to the Recall rate of the model, while there is no significant change between the model with other features and the baseline model. For the size and governance features, combined with the significant improvement in the Precision rate, the results show that these two types of features are more useful in reducing the misidentification rate of non-purchase samples. At the same time, from the perspective of AUC, these two types of features have a significant incremental effect on improving the effect of the model. When we add all features of the audit object class to the baseline model, we observe significant differences in each evaluation index compared with the baseline model, and the group differences of most indicators reach the maximum, which verifies that most audit object features make significant incremental contributions to the model. Comparing the size and complexity features, we find that the latter have a larger incremental effect on the model, but this is not inconsistent with the feature importance explanation, as the effect of the model is often not a simple accumulation of the independent effects of a single feature. There may be complex interactions between features, which can affect the overall performance of the

Table 7
Significance of incremental effects of features.

Model	Mean	SD	T value	P value
Panel A: Precision				
Base	0.1361	0.0062		
Base + Size	0.1445	0.0063	16.3032	0.0000
Base + Complexity	0.1789	0.0091	39.8356	0.0000
Base + Governance	0.1570	0.0078	22.5412	0.0000
Base + Control	0.1366	0.0065	1.3071	0.2014
Base + Audit object	0.1997	0.0112	42.6585	0.0000
Panel B: Recall				
Base	0.7373	0.0281		
Base + Size	0.7374	0.0300	0.0344	0.9728
Base + Complexity	0.7898	0.0287	10.0741	0.0000
Base + Governance	0.7321	0.0261	−1.0798	0.2891
Base + Control	0.7373	0.0274	0.0240	0.9810
Base + Audit object	0.7800	0.0285	7.5460	0.0000
Panel C: AUC				
Base	0.7288	0.0121		
Base + Size	0.7383	0.0140	6.6231	0.0000
Base + Complexity	0.7865	0.0114	24.0064	0.0000
Base + Governance	0.7485	0.0110	8.6148	0.0000
Base + Control	0.7294	0.0123	0.5722	0.5716
Base + Audit object	0.7965	0.0135	25.5733	0.0000

Table 8
Permutation importance.

RF	GBDT	RUSB	LR	SVC	MLP
ABEM_Proxy	*ABEM_Proxy*	*Party_CG*	*Currentasset*	*A*	*Sale*
Big4	*Big4*	*ABEM_Proxy*	*Sale*	*Currentasset*	*Currentasset*
Absem_DACC	*Party_CG*	*SOE*	*A*	*NetA*	*ABEM_Proxy*
Current_ratio	*Institutenumber_sum*	*Big4*	*NetA*	*Big4*	*NetA*
HHI_B	*Is_Institute*	*Netinterest_ratio*	*ABEM_Proxy*	*Sale*	*Big4*
LongLia_A	*Absem_DACC*	*Sale*	*Big4*	*ABEM_Proxy*	*Absem_DACC*
Sale	*Law_index*	*Absem_DACC*	*Absem_DACC*	*Absem_DACC*	*Audittyp_score*
Party_CG	*Sale*	*Law_index*	*Audittyp_score*	*Audittyp_score*	*Sum_Readnum*
Manage_female	*Current_ratio*	*NetA*	*SOE*	*Sum_Readnum*	*A*
NetA	*INV_A*	*Currentasset*	*Sum_Readnum*	*Market*	*ShrHolder1*

model. Therefore, even if a feature is important, if its interaction with other features is not strong, its incremental effect on the model effect may be limited.

5.2. Robustness testing of feature importance

Because the importance of features in the main results is indirectly measured by the feature coefficient of the LR model, to test the robustness of our results, we also use other methods to measure the importance of features. We mainly use permutation importance and Shapley additive explanations[4] to obtain feature importance. The basic idea of permutation importance is to evaluate the importance of features by arranging features randomly and measuring the change in model performance. If a feature has an important impact on the recognition performance of the model, its random arrangement will significantly reduce the accuracy of the model. SHAP is based on the famous Shapley value theory in game theory. It infers the decision of the model by calculating the effective value of each feature, to obtain meaningful interpretation results.

5.2.1. Permutation importance

By calculating the permutation importance of features, we analyze the feature importance of each model and list the top 10 features with permutation importance in Table 8. The results show that the SVC and MLP models exhibit a high degree of coincidence with the features of the LR model. Nine of the top 10 features are the same, but the ranking is slightly different. At the same time, by comparing the ranking of the importance of features in the main results, we find that in addition to the features of the shareholding ratio of the largest shareholder (*ShrHolder1*), the importance of other features, including total assets (*A*), current assets (*Currentasset*), accounting firm size (*Big4*), operating revenue (*Sales*), real earnings management (*ABEM_Proxy*) and net assets (*NetA*) of the auditee, is verified. The RUSBoost model and the LR model are also consistent in identifying the importance of features. Seven of the top 10 important features are the same as the LR model. The RF and GBDT models differ greatly from the LR model in identifying the importance of features, with only five and four overlapping features, respectively. However, we can observe that the importance of the first 10 features of the LR model is verified in other models, which further confirms the LR model's judgment on feature importance.

5.2.2. SHAP value

We also obtain the feature importance ranking based on the SHAP value. We analyze the LR and SVC models and the results are shown in Fig. 1 and Fig. 2, respectively. Fig. 1 presents the results for the LR model. It can be seen that the eight most important features coincide with those in the main results, but the ranking is slightly different. The importance of features, including the shareholding ratio of the largest shareholder (*ShrHolder1*), is verified. Fig. 2 presents the results for the SVC model. The eight most important

[4] This is implemented using the Eli5 and Shap libraries in Python.

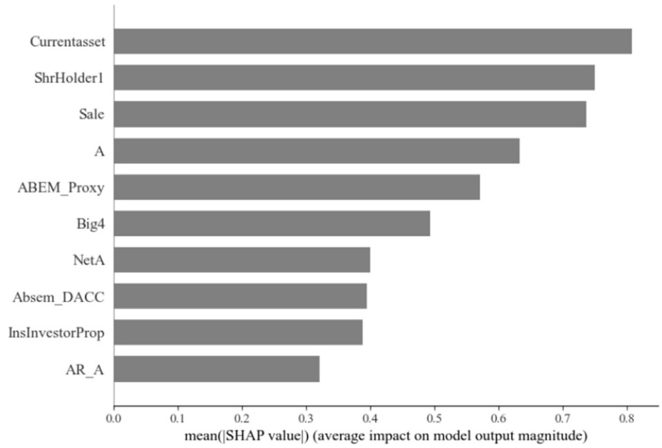


Fig. 1. Feature importance based on LR model (SHAP value).

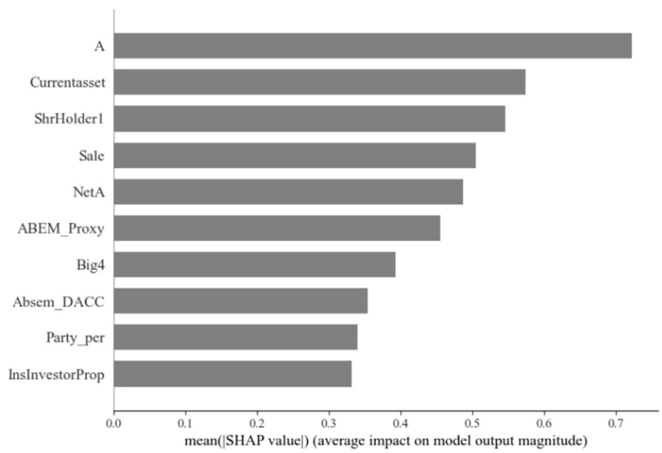


Fig. 2. Feature importance based on SVC model (SHAP value).

features are consistent with those of the LR model, but there are some differences in the ranking of total assets (*A*) and net assets (*NetA*) of the audited entity between the two models. Therefore, these results confirm the six most important features for identifying audit opinion shopping behavior analyzed in the main results: accounting firm size (*Big4*), current assets (*Currentasset*), operating revenue (*Sales*), total assets (*A*), shareholding ratio of the largest shareholder (*ShrHolder1*) and real earnings management of the audited entity (*ABEM_Proxy*).

5.3. Correlation between features

To prevent the negative effect of potential multicollinearity between feature variables on the recognition effect of the model, we conduct a collinearity test on the feature variables. Specifically, we use the variance inflation factor (VIF) method to strictly evaluate the collinearity of the selected feature variables. As a quantitative index of test results, the VIF value can directly reflect the degree of collinearity between variables. In general, when the VIF value is between 0 and 10, there is no significant collinearity problem between the variables. Table 9 presents the VIF values of each feature variable, excluding the variable *Province*, indicating where the office is located. According to the results in Table 9, the VIF values of six variables, namely total

Table 9
VIF values for each feature.

No.	Feature	VIF	No.	Feature	VIF
1	*A*	59.96	30	*Litigation_Risk*	1.11
2	*Sale*	7.31	31	*DeficiencyType*	1.17
3	*Netinterest_ratio*	2.16	32	*IC*	1.96
4	*NetA*	16.09	33	*Big4*	1.21
5	*Currentasset*	24.55	34	*BETA1Year2*	1.38
6	*AR_A*	1.75	35	*STPT*	1.34
7	*INV_A*	1.86	36	*LongLia_A*	1.90
8	*T5cusoirt*	1.23	37	*Current_ratio*	2.03
9	*T5suplpart*	1.21	38	*Market*	4.56
10	*Absem_DACC*	1.13	39	*HHI_B*	1.04
11	*ABEM_Proxy*	1.18	40	*Audittyp_score*	1.44
12	*Restate*	1.05	41	*Law_index*	5.91
13	*Restate0*	1.04	42	*J_F_News*	1.16
14	*InDrcRat*	1.10	43	*J_F_Online*	1.39
15	*ShrHolder1*	1.64	44	*Investor_neg*	1.21
16	*Pld_dum*	1.37	45	*Investor_pos*	1.40
17	*Party_per*	19.43	46	*Sum_Readnum*	5.43
18	*Party_per_IsMTMT*	4.67	47	*Sum_Commentnum*	4.75
19	*Party_per_IsMTB*	7.76	48	*AnaAttention*	39.65
20	*Party_per_IsSupervisor*	3.23	49	*InsInvestorProp*	2.46
21	*Party_CG*	1.87	50	*ReportAttention*	40.05
22	*GmShrRat*	1.59	51	*Is_Institute*	3.70
23	*SOE*	2.47	52	*Institutenumber*	6.20
24	*ESO0*	1.05	53	*Institutenumber_sum*	6.43
25	*ESO*	1.07	54	*Hubris*	1.17
26	*IsAuditCommitteeSetUp*	1.03	55	*Manage_female*	1.08
27	*Loss*	2.02	56	*AreaR*	1.14
28	*Firm_age*	1.19	57	*MSR*	1.14
29	*Financial_Risk*	2.51			

assets (*A*), net assets (*NetA*), current assets (*Currentasset*), total proportion of Party members among directors, supervisors and senior management (*Party_per*), analyst attention (*AnaAttention*) and research report attention (*ReportAttention*), are greater than 10, so feature removal is necessary. As total assets are strongly correlated with net assets and current assets, net assets and current assets are excluded and a broader indicator of total assets is used. There is an inclusion relationship between the total proportion of Party members among directors, supervisors, and senior management and other feature variables in the same category, so this feature is eliminated. There is a strong correlation between the number of analysts following a company and the number of research reports. Ultimately, we retain the number of analysts. Finally, four feature variables, namely net assets (*NetA*), current assets (*Currentasset*), total proportion of Party members among directors, supervisors and senior management (*Party_per*) and research report attention (*ReportAttention*), are eliminated and 54 features are retained. The VIF test is performed again and the results show that the VIF value of each feature is less than 10, meeting the requirements of the collinearity test and proving that there is no serious multicollinearity between these feature variables.

After eliminating the above features, we retrain and retest the LR model. By analyzing the confusion matrix of the new test results, we find that out of the 232 target samples in the test set, 177 samples with audit opinion shopping can be accurately identified. Compared with the results before removing these features, the real class and true negative class are slightly reduced, indicating that removing features will weaken the recognition effect of the model. Additionally, we obtain a new feature importance and list the top six features in the feature importance ranking in Table 10. The results show that after removing current assets and net assets, the results of the other eight features in the top 10 features remain, which verifies our analysis of feature importance.

Table 10
Feature importance removing some features.

Feature	Category	Subcategory	Feature importance
Big4	Audit subject	Size	20.22 %
ShrHolder1	Audit object	Governance	10.94 %
ABEM_Proxy	Audit object	Complexity	10.51 %
Sale	Audit object	Size	9.75 %
A	Audit object	Size	6.40 %
Absem_DACC	Audit object	Complexity	5.14 %
AR_A	Audit object	Complexity	5.06 %
Audittyp_score	Audit subject	Audit result	4.21 %
Current_ratio	Audit subject	Risk premium	4.11 %
InsInvestorProp	Audit environment	Monitoring environment	3.36 %

6. Further analysis

6.1. Interpretability analysis

Feature importance can only help analyze which features are important for model recognition, but how these features affect the output of the model is unknown. Therefore, we use the Beeswarm plot in the SHAP library to visualize the contribution of features to the model recognition results. Fig. 3 is a Beeswarm plot drawn for the LR model identification process on the test set, which can be used to explain the model comprehensively. Each dot on the graph represents a sample. The closer the color is to red, the higher the value of the feature; the closer the color is to blue, the lower the value. If the feature is blue on the left and red on the right, this indicates that this feature contributes positively to the probability of recognition as audit opinion shopping, and vice versa. From Fig. 3, the shareholding ratio of the largest shareholder (*ShrHolder1*), operating revenue (*Sales*), total assets (*A*) and accrual-based earnings management (*Absem_DACC*) are positive features. The higher the value, the higher the probability that the LR model identifies a sample as an audit opinion shopping firm. Current assets (*Currentasset*), real earnings management (*ABEM_Proxy*), net assets (*NetA*) and accounting firm size (*Big4*) are negative features. The higher the values of these features, the lower the probability that the sample will be identified as an audit opinion shopping firm. According to the interpretability analysis, the larger the total size of a firm's assets, the higher the probability of being identified as purchasing audit opinions. Larger firms may exert greater performance pressure on auditors, which is more

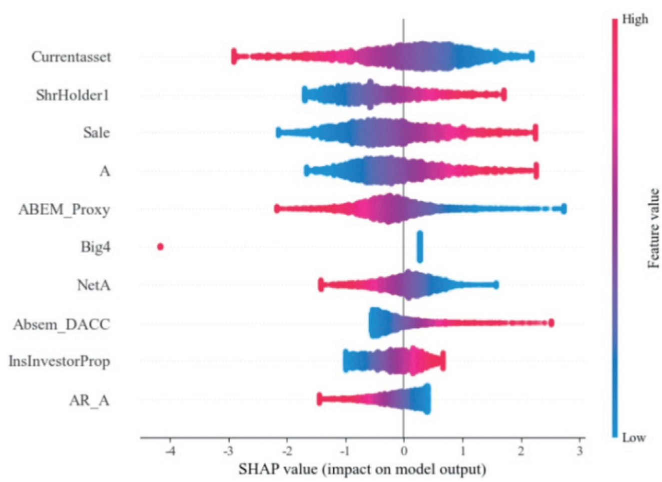


Fig. 3. Bee swarm based on LR model.

Table 11
Permutation importance based on different labels.

Ranking	*Purchase_sum*	*Purchase_change*	*Purchase_abfee*
1	*Currentasset*	*Currentasset*	*Currentasset*
2	*Sale*	*Sale*	*Sum_Commentnum*
3	*A*	*Big4*	*AnaAttention*
4	*NetA*	*NetA*	*Sum_Readnum*
5	*ABEM_Proxy*	*ABEM_Proxy*	*A*
6	*Big4*	*A*	*Party_per_IsSupervisor*
7	*Absem_DACC*	*Absem_DACC*	*Netinterest_ratio*
8	*Audittyp_score*	*SOE*	*ABEM_Proxy*
9	*SOE*	*Audittyp_score*	*Current_ratio*
10	*Sum_Readnum*	*Party_per_IsMTB*	*AreaR*

important than auditors, thus strengthening collusion between the two parties (Fang et al., 2020). The shareholding ratio of the largest shareholder is a positive feature, which is consistent with previous studies. When the shareholding ratio of the largest shareholder is too high, this shareholder will use their own equity advantage to engage in tunneling operations and tend to collude with the firm (Francis and Wilson, 1988; Li and Ren, 2012). Operating revenue is a positive feature, which may indicate the existence of sales manipulation or early recognition of revenue and the purchase of audit opinions to hide. As for the positive contribution of accrual-based earnings management, Chen Jun et al. (2021) find that earnings management manipulation strengthens listed companies' motivation to engage in audit opinion shopping, while Cao et al. (2013) argue that firms with more accrual-based earnings management choose to purchase audit opinions to avoid being issued non-standard audit opinions. Regarding the negative features, first, accounting firm size is relatively easy to understand. Previous studies show that accounting firm size is positively correlated with audit quality (Wu et al., 2015), with the "Big Four" providing better audit quality (Xu and Qi, 2023). Therefore, the larger the size of the accounting firm, the higher its independence and the lower the probability of collusion. The negative effect of real earnings management on model identification may be due to the fact that regulators and auditors pay less attention to real earnings management than high-risk accrual-based earnings management (Cohen et al., 2008), which is relatively hidden (Xie et al., 2013). Moreover, due to the mutual substitution relationship between the two earnings management methods, a low degree of real earnings management is accompanied by a high degree of accrual-based earnings management, so the lower the degree of real earnings management, the higher the probability that the firm will be identified as having purchased audit opinions by the model. The negative contribution of net assets and current assets may be due to the fact that when firms face lower operational and financial risks, they have less incentive to purchase audit opinions.

6.2. Feature importance of two types of audit opinion shopping

We also analyze different ways to purchase audit opinions, classify and count the permutation importance of features based on the LR model and list the top 10 permutation importance features in Table 11. The results

Table 12
Model evaluation adding cross features.

Model	Precision	Recall	AUC
Univariate features	0.2096	0.7716	0.7938
Univariate features + 21 cross features	0.2106	0.7716	0.7944
Univariate features + 5 cross features	0.2144	0.7716	0.7964
Univariate features + *ShrHolder1_InsInvestorProp*	0.2156	0.7845	0.8021
Univariate features + *Audittyp_score_Sale*	0.2094	0.7716	0.7937
Univariate features + *Big4_ShrHolder1*	0.2098	0.7716	0.7940
Univariate features + *Audittyp_score_Absem_DACC*	0.2065	0.7629	0.7888
Univariate features + *Big4_Absem_DACC*	0.2098	0.7716	0.7940
Univariate features + 3 cross features	0.2162	0.7845	0.8023

Table 13
Feature importance adding 3 cross features.

Feature	Feature importance
ShrHolder1_InsInvestorProp	19.06 %
Currentasset	11.72 %
Sale	11.47 %
Big4_ShrHolder1	9.64 %
ABEM_Proxy	7.80 %
A	6.26 %
NetA	3.79 %
Absem_DACC	3.68 %
Audittyp_score	2.93 %
Big4_Absem_DACC	2.91 %

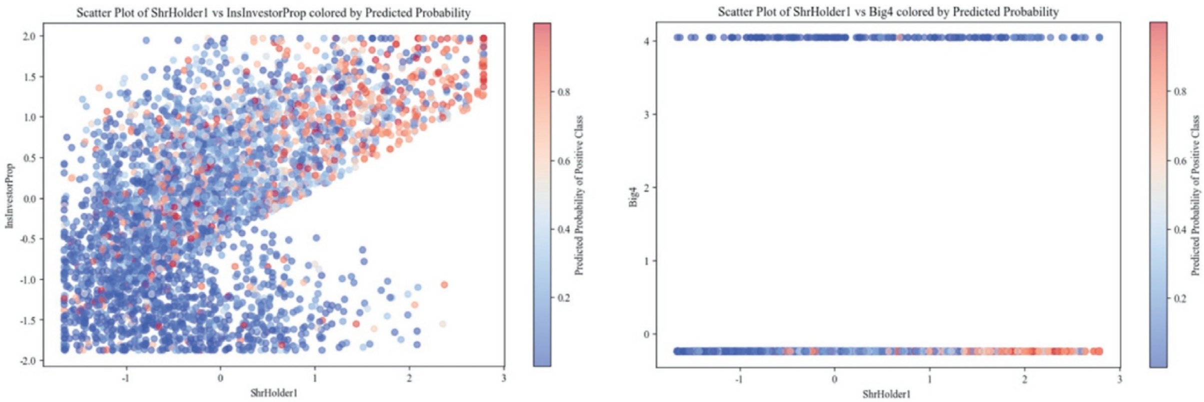
show that the identification of feature importance for changing audit firms is basically consistent with the consolidation model, but the importance of firm size is higher. When modeling the purchase of audit opinions using paying abnormal audit fees, the model also pays attention to the two features of current assets (*Currentasset*) and total assets (*A*), but other features change. The features of the monitoring environment, such as *Sum_Commentnum*, *Sum_Readnum* and *AnaAttention*, and governance features, such as *Party_per_IsSupervisor*, are more important and are related to supervision. This finding is understandable, because in a strong supervision environment, firms may choose a more covert way to purchase audit opinions by paying abnormal audit fees.

6.3. Feature crossover

Unlike general corporate violations, purchasing audit opinions also involves interaction between stakeholders. Therefore, in addition to univariate features, combined features can affect the effect of the model. We use a total of 58 feature variables across three categories. Pairwise intersection of feature variables from different categories produces 975 feature combinations, which may lead to a risk of overfitting and "dimensional disaster." Therefore, we only cross the important features and add the cross features to the model for training and testing. The test results are presented in Table 12. Specifically, we perform a pairwise crossover of the top 10 feature categories evaluated by the LR model. This mainly involves crossing the audit object with the other two categories, resulting in 21 cross features.[5] Twenty-one cross features are added to the feature variables for model training and testing. The results show that the Precision and AUC values of the model improve compared with the univariate feature model, indicating that the recognition effect of the model is better. In addition, we screen the top five cross features, namely *ShrHolder1_InsInvestorProp*, *Audittyp_score_Sale*, *Big4_ShrHolder1*, *Audittyp_score_Absem_DACC* and *Big4_Absem_DACC*, using feature importance to identify the most effective feature combination. By adding these five cross feature variables to the model separately, we find that the three cross features of *ShrHolder1_InsInvestorProp*, *Big4_ShrHolder1* and *Big4_Absem_DACC* play a positive role in improving the effect of the model. Finally, we add these three cross features to the model and find that the AUC value of the model on the test set reaches 0.8023, while the Precision and Recall rates are also higher than those of other models.

Furthermore, we analyze the feature importance of the three added cross features in the LR model and list the top 10 features in the feature importance ranking in Table 13. The results show that the cross feature of the shareholding ratio of the largest shareholder and institutional investors (*ShrHolder1_InsInvestorProp*) is the most important, indicating that the interaction between major shareholders and institutional investors will seriously affect the audit opinion shopping behavior of firms. When the shareholding ratio of major sharehold-

[5] We only cross the features of the audit object with the other two feature categories because, as there is no audit object, audit opinion shopping behavior will lack the behavior subject, and the features of the audit object and the audit subject, as well as the features of the audit object and the audit environment will be crossed, resulting in 21 cross features.



Panel A: *ShrHolder1* and *InsInvestorProp* Panel B: *Big4* and *ShrHolder1*

Fig. 4. Feature scatter plot.

ers is too high, on the one hand, institutional investors play a governance and supervisory role, which may inhibit the tunneling behavior of major shareholders and thus the purchase of audit opinions. On the other hand, institutional investors may engage in "distraction" or "group" behavior, resulting in governance and supervision failure, and even conspire with major shareholders to intensify their tunneling behavior, thus promoting the purchase of audit opinions (Wang et al., 2022). The other two cross features, accounting firm size and shareholding ratio of the largest shareholder (*Big4_ShrHolder1*) and accounting firm size and degree of accrual-based earnings management (*Big4_Absem_DACC*), are also of great importance to the model, both of which are related to accounting firm size. When an accounting firm is small and relatively independent, it will be more convenient for large shareholders to encroach on their interests and for firms to conceal earnings management, which may promote the purchase of audit opinions.

Because we use multiplication as the method of feature intersection, it is difficult to intuitively observe the interaction between features. Therefore, we combine the scatter diagram for analysis. Specifically, based on the univariate LR model, we draw the scatter diagram between the two cross features and the model recognition results. As shown in Fig. 4, each dot represents a pair of observation values in the test set, in which the value of one feature is represented on the x-axis and the value of the other feature is represented on the y-axis, and the color of the dot represents the recognition probability. The closer the color is to red, the greater the probability that the sample will be identified as having purchased audit opinions. As can be seen from Panel A in Fig. 4, the red dots are more concentrated in the upper right corner, indicating that when the shareholding ratio of institutions and the shareholding ratio of the largest shareholder are both high, firms are more likely to purchase audit opinions. In particular, when the shareholding ratio of institutions is the highest, with the increase in the shareholding ratio of the largest shareholder, the dots change from blue to red from left to right, indicating that with the increase in the shareholding ratio of the largest shareholder, the supervisory role of institutional investors gradually weakens. When the shareholding ratio of the largest shareholder reaches its maximum, the change in the institutional shareholding ratio will hardly affect the change in identification probability. Similarly, in Panel B, the distribution of the dots represents two straight lines because we use a dummy variable to measure accounting firm size. The results show that when a firm is a "Big Four" accounting firm, the change in the shareholding ratio of the largest shareholders has no significant impact on the probability of recognition, while when the firm is not a "Big Four" accounting firm, with the increase in the shareholding ratio of the largest shareholder, the probability that the sample will be identified as having purchased audit opinions increases. This shows that poor accounting firm independence makes it easier for major shareholders to engage in tunneling behavior, making it easier for firms to purchase audit opinions.

Table 14
Model evaluation based on the prediction.

Model	Precision	Recall	AUC	Precision@120	Recall@120	NDCG@120
LR	0.1881	0.7967	0.8027	0.4333	0.4228	0.4522
SVC	0.1712	0.7724	0.7821	0.4167	0.4065	0.4417

Table 15
Feature importance based on the prediction.

Feature	Category	Subcategory	Feature importance
Big4	Audit subject	Size	28.44 %
Audittyp_score	Audit subject	Audit result	24.10 %
ShrHolder1	Audit object	Governance	10.24 %
InsInvestorProp	Audit environment	Monitoring environment	8.85 %
Sale	Audit object	Size	6.66 %
ABEM_Proxy	Audit object	Complexity	2.42 %
AR_A	Audit object	Complexity	2.34 %
Current_ratio	Audit subject	Risk premium	2.20 %
NetA	Audit object	Size	2.03 %
Currentasset	Audit object	Size	1.96 %

6.4. Forecasts

Finally, we use a model to predict audit opinion shopping behavior. Specifically, the tag value of *Purchase_sum* is replaced with the next year's tag value. As our sample period ends in 2021, data for 2022 are missing, so the sample for 2021 is dropped. The dataset is divided at a ratio of 7:3. The training set is used for model learning and parameter optimization, while the test set is used to evaluate the generalization ability of the model on unknown data. Next, we use the training set to train the model systematically and make predictions on the test set. Using the confusion matrix to analyze the prediction results, the matrix shows that 98 of the 123 target samples are accurately predicted. We also analyze the performance of the LR and SVC models on the test set and the evaluation indices are presented in Table 14. We find that the models also have a good predictive effect on audit opinion shopping behavior in the following year. The Recall rate of the LR model is 0.7967, indicating that the model can predict 79.67 % of the samples that will purchase audit opinions, and the AUC value reaches 0.8027, indicating that the model has a good classification effect in general. However, the predictive effect of the LR model is slightly better than that of the SVC model.

Next, we analyze the feature importance of the LR prediction model on the test set and the first 10 important features are presented in Table 15. We see some changes in the importance of features in the prediction model. The importance of audit results (*Audittyp_score*) and the shareholding ratio of institutional investors (*InsInvestorProp*) in this year increases, indicating that when predicting the purchase behavior of audit opinions, we should pay more attention to audit results and institutional shares in the year preceding the prediction year. At the same time, the importance of accounts receivable (*AR_R*) and current ratio (*Current_ratio*) increases, indicating that business complexity and liquidity risk are important factors to consider when making audit opinion shopping decisions.

7. Conclusions

Auditing serves as an assurance mechanism for firms' financial reports, with audit opinions representing auditors' evaluation of their authenticity and compliance. However, the phenomenon of audit opinion shopping often occurs in practice. Audit opinions act as information intermediaries for stakeholders assessing listed firms, and low-quality opinions can disrupt information symmetry, resulting in losses for non-listed stakeholders and harming capital market efficiency. Thus, it is essential to establish a scientific early warning system against audit opinion shopping to maintain audit quality, protect investor interests and promote market fairness and transparency.

This study uses machine learning to identify out-of-sample audit opinion shopping behavior, with the aim of obtaining scientific early warning of audit risks. Among the six models tested, the LR model is selected due to its superior classification performance and faster operational speed. The LR model is then used to derive importance scores for each feature. Key features for identifying audit opinion shopping include firm size, current assets, operating revenue, total assets, ownership of the largest shareholder and the degree of real earnings management. The audit object emerges as the most critical feature, with most object-related features significantly enhancing the model's classification capability.

Our robustness tests verify the model's applicability, using alternative calculation methods to assess feature importance. Further analysis includes model visualizations and exploration of feature interactions through cross features. A predictive model for audit opinion shopping is also developed.

This research has two important aspects. The first is theoretical. This study proposes the integration of machine learning algorithms to identify audit opinion shopping, thus offering a novel approach to early warning of audit collusion risks and serving as a valuable addition to current research in this field. Second, the practicality of the model is evident in its ability to facilitate the identification of audit collusion behavior in a more scientific and reasonable manner. This, in turn, leads to improved audit quality and protection of investors' interests. Furthermore, it enhances market transparency by revealing potential misconduct and providing investors with more accurate information to mitigate risks. This, in turn, has the effect of improving market governance and promoting capital market efficiency.

However, our study has some limitations. Data access constraints lead to incomplete selection of feature variables, potentially affecting our research results. Moreover, although the Chinese context offers unique data availability and policy insights, the applicability of the findings to other economies remains uncertain due to the diversity of economic, market and regulatory landscapes. Future research should expand data sources, incorporate more relevant variables and adapt models to different economic contexts to enhance their explanatory power and universality. Additionally, more direct measurement methods or combined data sources could be employed to better capture the real situation of label variables.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the editor and referee for their constructive and valuable comments. We acknowledge financial support from the National Natural Science Foundation of China (72372160) and the Fundamental Research Funds for the Central Universities at the Zhongnan University of Economics and Law (2722024BY016).

Appendix A. Variable definitions

Feature	Definition
A	Total assets at the end of the year, computed as the natural logarithm of (Total assets + 1).
Sale	Total operating revenue at the end of the year, calculated as the natural logarithm of (Operating revenue + 1).
Netinterest_ratio	Net profit margin, equal to (Net profit / Operating revenue).
NetA	Net assets at the end of the year, calculated as the natural logarithm of (Total owner's equity + 1).
Currentasset	Current assets at the end of the year, calculated as the natural logarithm of (Total current assets + 1).
AR_A	Accounts receivable ratio, equal to (Accounts receivable / Total assets).
INV_A	Inventory ratio, equal to (Inventory / Total assets).
T5cusoirt	Revenue proportion from top five customers.
T5suplpart	Purchase proportion from top five suppliers.
Absem_DACC	Absolute value of accrual-based earnings management, measured using the Jones model. A higher |DACC| indicates a higher degree of earnings management.
ABEM_Proxy	Real earnings management level, measured from three aspects: production manipulation, sales manipulation, and discretionary expense manipulation. Real earnings management = Production manipulation − Sales manipulation − Discretionary expense manipulation.
Restate	A dummy variable that equals 1 if financial statements have been restated, and 0 otherwise.
Restate0	A dummy variable that equals 1 if financial statements were restated in the previous year, and 0 otherwise.
InDrcRat	Independent director ratio.
ShrHolder1	Shareholding ratio of the largest shareholder.
Pld_dum	A dummy variable that equals 1 if there is equity pledge by the controlling shareholder, and 0 otherwise.
Party_per	Total proportion of Party members among directors, supervisors, and senior management.
Party_per_IsMTMT	Proportion of Party members among senior management.
Party_per_IsMTB	Proportion of Party members among the board of directors.
Party_per_IsSupervisor	Proportion of Party members among the supervisory board.
Party_CG	A dummy variable that equals 1 if the Party organization participates in corporate governance, and 0 otherwise. The value is 1 if the Party secretary concurrently serves as the chairman, supervisory board chairman, or general manager; otherwise, it is 0.
GmShrRat	Shareholding ratio of senior management.
Manage_female	Proportion of female senior management.
Hubris	Overconfidence of senior management, measured as (Total compensation of top three senior managers / Total compensation of senior managers).
SOE	A dummy variable that equals 1 if the firm is a state-owned enterprise, and 0 otherwise.
ESO	A dummy variable that equals 1 if an equity incentive plan is announced in the current year, and 0 otherwise.
ESO0	A dummy variable that equals 1 if an equity incentive plan will be implemented in the following year, and 0 otherwise.

(*continued on next page*)

Appendix A (*continued*)

Feature	Definition
IsAuditCommitteeSetUp	A dummy variable that equals 1 if an audit committee has been established, and 0 otherwise.
Loss	A dummy variable that equals 1 if the firm incurs losses, and 0 otherwise.
Province	Province where the office is located.
Firm_age	Year of establishment of the company, calculated as Ln(Year of establishment).
Financial_Risk	Financial risk, measured using the Z-score, with values ranging from 1 to 3. A lower value indicates higher risk.
Litigation_Risk	A dummy variable that equals 1 if the firm has any litigation cases in the current year, and 0 otherwise.
IC	Internal control index / 100, sourced from the Dibo database.
Deficiency_Type	Type of internal control deficiency, 0 = No deficiency; 1 = Material deficiency; 2 = Significant deficiency; 3 = General deficiency.
Big4	A dummy variable that equals 1 if the accounting firm is one of the Big Four accounting firms, and 0 otherwise.
Beta	BETA value, weighted by total market value for one year, measuring the operational risk of the inspected unit.
Stpt	A dummy variable that equals 1 if the firm is flagged as ST, *ST, or PT, and 0 otherwise.
LongLia_A	Asset long-term liability ratio, equal to (Long-term liabilities / Total assets).
Current_ratio	Current ratio, equal to (Current assets / Current liabilities).
Market	Marketization index, the total index from the marketization index report.
HHI	Sum of the squares of the ratio of each company's total owner's equity to the industry's total owner's equity.
AreaR	Regional expertise of the accounting firm, calculated using the regional market share method based on the audit revenue of listed companies.
MSR	Industry expertise of the accounting firm, calculated using the industry market share method based on the audit revenue of listed companies.
Audittyp_score	Audit opinion, ranked from strict to lenient: Disclaimer opinion or adverse opinion; Qualified opinion; Unqualified opinion with emphasis-of-matter paragraph; Standard unqualified opinion. Corresponding values range from 4 to 1.
Law_index	Index of the development of market intermediary organizations and the legal institutional environment.
J_F_News	Press media supervision pressure, measured using the Janis-Fadner coefficient, with values ranging from −1 to 1. A value closer to 1 indicates less press media supervision pressure on the company; a value closer to −1 indicates greater supervision pressure.
J_F_Online	Online media supervision pressure, measured using the Janis-Fadner coefficient, with values ranging from −1 to 1. A value closer to 1 indicates less online media supervision pressure on the company; a value closer to −1 indicates greater supervision pressure.
Investor_neg	Proportion of negative investor posts.
Investor_pos	Proportion of positive investor posts.
Sum_Readnum	Reading volume, calculated as the natural logarithm of (1 + Total reading volume of posts in the year).
Sum_Commentnum	Comment volume, calculated as the natural logarithm of (1 + Total comment volume of posts in the year).

Appendix A (*continued*)

Feature	Definition
AnaAttention	Analyst attention, the number of analyst teams that have tracked and analyzed the company within one year, with the final result plus one and then taking the logarithm.
ReportAttention	Research report attention, the number of research reports that have tracked and analyzed the company within one year, with the final result plus one and then taking the logarithm.
InsInvestorProp	Shareholding ratio of institutional investors.
Institutenumber	Number of institutions involved in the survey, calculated as the natural logarithm of (1 + Number of surveyed institutions).
Institutenumber_sum	Frequency of surveys, calculated as the natural logarithm of (1 + Total number of surveys).
Is_Institute	A dummy variable that equals 1 if institutional investors conduct a survey, and 0 otherwise.

Appendix B. Model evaluation using alternative splits of data set

Model	Precision	Recall	AUC	Precision@k	Recall@k	NDCG@k
Panel A: Split the sample according to the ratio 8:2						
RF	0.4773	0.2675	0.6243	0.4667	0.4459	0.4976
GBDT	0.5281	0.2994	0.6411	0.4000	0.3822	0.4467
RUSB	0.3926	0.3376	0.6520	0.3867	0.3694	0.3826
LR	0.2095	0.7580	0.7871	0.4000	0.3822	0.4181
SVC	0.2063	0.7962	0.7997	0.4133	0.3949	0.4311
MLP	0.4000	0.3822	0.6727	0.4000	0.3822	0.4418
Panel B: Split the sample by year						
RF	0.4023	0.1707	0.5771	0.3700	0.3610	0.3929
GBDT	0.4104	0.2683	0.6216	0.3550	0.3463	0.3992
RUSB	0.3557	0.3366	0.6485	0.3600	0.3512	0.3987
LR	0.1912	0.7610	0.7758	0.4100	0.4000	0.4450
SVC	0.1864	0.7902	0.7830	0.4450	0.4341	0.4728
MLP	0.3438	0.4293	0.6880	0.3750	0.3659	0.4010

Note: (1) In this table, we have made changes to the division of testing and training sets. Firstly, we divided the data set by year, taking data from 2014 to 2019 as the training set and data from 2020 to 2021 as the test set. At the same time, we divided the data set according to the ratio of 8: 2 for the whole sample. (2) The parameters of each model are the same as in the previous section. (3) In Panel A, the number of samples in the test set is 2601, and when rounded according to the proportion k of the target samples, it is taken as 150; in Panel B, the number of samples in the test set is 3358, and when rounded according to the proportion k of the target samples, it is taken as 200.

Appendix C. Model evaluation using alternative sampling methods

Model	Precision	Recall	AUC	Precision@k	Recall@k	NDCG@k
Panel A: Random undersampling						
RF	0.1959	0.8190	0.8032	0.3913	0.3879	0.3924
GBDT	0.2037	0.8147	0.8066	0.3696	0.3664	0.3742
RUSB	0.2000	0.8060	0.8011	0.3609	0.3578	0.3733
LR	0.2008	0.8233	0.8081	0.3826	0.3793	0.3908
SVC	0.2114	0.8017	0.8063	0.4000	0.3966	0.4042
MLP	0.1874	0.8233	0.7988	0.3652	0.3621	0.3769
Panel B: SMOTEENN						
RF	0.3134	0.4741	0.7042	0.3522	0.3491	0.3822
GBDT	0.3469	0.5129	0.7259	0.3957	0.3922	0.4065
RUSB	0.2394	0.5603	0.7239	0.3391	0.3362	0.3349
LR	0.1664	0.8319	0.7842	0.3783	0.3750	0.3834
SVC	0.1801	0.8103	0.7886	0.3957	0.3922	0.4033
MLP	0.2887	0.6034	0.7547	0.4000	0.3966	0.4104
Panel C: Class weighting						
RF	0.5647	0.2069	0.5984	0.4348	0.4310	0.4578
GBDT	0.2799	0.6853	0.7870	0.4261	0.4224	0.4548
RUSB	0.2287	0.7414	0.7917	0.3957	0.3922	0.4040
LR	0.2052	0.7759	0.7930	0.3870	0.3836	0.4019
SVC	0.1976	0.7716	0.7867	0.3913	0.3879	0.4102
MLP	0.5474	0.2241	0.6062	0.4087	0.4052	0.4499

Note: (1) The results in this table are that the methods of random undersampling, SMOTEENN mixed sampling and setting category weights are respectively used to deal with the problem of imbalanced samples. (2) Panel A is under random sampling, and the model parameters: max_features, min_samples_leaf, max_depth, min_samples_split and n_estimators of RF model are 13, 2, 17, 2 and 264 respectively; The learning_rate, n_estimators, max_depth, min_samples_leave of GBDT model are 0.27, 177, 1 and 12, respectively; The learning_rate and n_estimators of RUSBoost model are 0.4 and 90, respectively; The solver, penalty and C of LR model are respectively "liblinear", "l1" and "1"; The C, gamma and kernel of SVM are 64, 0.001 and "rbf" respectively; The hidden_layer_sizes, max_iter, activation, alpha, solver of MLP model are (400,), 20, "relu", 0.0001, "adam", respectively. (3) Panel B is SMOTEENN sampling, and the model parameters: max_features, min_samples_leaf, max_depth, min_samples_split and n_estimators of RF model are "log2", 2, 26, 2 and 240 respectively; The learning_rate, n_estimators, max_depth, min_samples_leave of GBDT model are 0.12, 168, None, and 14 respectively; The learning_rate and n_estimators of RUSBoost model are 1 and 270 respectively; The solver, penalty and C of LR model are "linear", "l1" and 1 respectively; The C, gamma and kernel of SVM are 100, 0.1 and "linear" respectively; The hidden_layer_sizes, learning_rate, max_iter, activation, alpha, learning_rate_init, solver And batch_size of MLP model are (220,), "constant", 80, "logistic", 0.001, 0.01, "adam" and 90, respectively. (4) In Panel C, to assign weights to the minority class, the parameters for each model are set as follows: for the RF model, the parameters class_weight, max_features, min_samples_leaf, max_depth, min_samples_split, and n_estimators are set to "balanced", 15, 5, 20, 2, and 900, respectively;

for the GBDT model, the parameters learning_rate, n_estimators, max_depth, and min_samples_leaf are set to 0.1, 150, 3, and 40, respectively; for the RUSBoost model, the parameters learning_rate and n_estimators are set to 0.1 and 239, respectively; for the LR model, the parameters class_weight, solver, penalty, and C are set to "balanced", "linear", "l1", and 0.1, respectively; for the SVM model, the parameters class_weight, C, gamma, and kernel are set to "balanced", 100, 0.1, and "linear", respectively; for the MLP model, the parameters hidden_layer_sizes, learning_rate, max_iter, activation, alpha, learning_rate_init, solver, and batch_size are set to (200,), "adaptive", 280, "tanh", 0.001, 0.001, "sgd", and 100, respectively. (5) The value of k in the table is set to 230.

References

Athey, S., Imbens, G.W., 2019. Machine learning methods that economists should know about. Annu. Rev. Econ. 11 (1), 685–725.

Bao, Y., Ke, B., Li, B., 2020. Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. J. Account. Res. 58 (1), 199–235.

Bertomeu, J., Cheynel, E., Floyd, E., Pan, W., 2021. Using machine learning to detect misstatements. Rev. Acc. Stud. 26 (2), 468–519.

Brown, N.C., Crowley, R.M., Elliott, W.B., 2020. What are you saying? Using topic to detect financial misreporting. J. Account. Res. 58 (1), 237–291.

Cao, Q., Bu, H., Yang, Y., Liu, C., 2013. Earnings management, audit fee and audit opinions. Audit. Res. 6, 76–83 (in Chinese).

Cao, F., Li, K., 2019. Controlling shareholder's share pledge and audit opinion shopping of listed companies. Audit. Res. 2, 108–118 (in Chinese).

Cohen, D.A., Aiyesha, D., Thomas, Z.L., 2008. Real and accrual-based earnings management in the pre- and post-sarbanes-oxley periods. Account. Rev. 83 (3), 757–787.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357.

Chen, Y., Gul, F.A., Veeraraghavan, M., Zolotoy, L., 2015. Executive equity risk-taking incentives and audit pricing. Account. Rev. 90 (6), 2205–2234.

Chen, F., Peng, S., Xue, S., Ye, F., 2016. Do audit clients successfully engage in opinion shopping? partner-level evidence. J. Artif. Intell. Res. 54 (1), 79–112.

Chen, S., Cao, Y., 2018. Audit opinion shopping and executive stock option incentive–based on the perspective of cheap talk game. Audit. Res. 1, 59–67 (in Chinese).

Chen, J., Xu, H., Lin, J., 2021. How does enterprise rent-seeking affect audit opinion shopping?. Account. Res. 7, 180–192 (in Chinese).

Chen, L., Sun, M., Ji, N., Zhang, L., 2023. Will the change of auditors affect the textual similarity in key audit matters?. Audit. Res. 3, 72–84 (in Chinese).

Choudhury, P., Wang, D., Carlson, N.A., Khanna, T., 2019. Machine learning approaches to facial and text analysis: discovering CEO oral communication styles. Strateg. Manag. J. 40 (11), 1705–1732.

DeFond, M.L., Subramanyam, K.R., 1998. Auditor changes and discretionary accruals. J. Account. Econ. 25 (1), 35–67.

DeFond, M., Zhang, J., 2014. A review of archival auditing research. J. Account. Econ. 58 (2–3), 275–326.

Du, X., Guo, J., 2008. Auditor switching and audit opinion-shopping: an empirical research. J. Shanxi Univ. Financ. Econ. 30 (11), 101–106 (in Chinese).

Fan, S., Kong, D., Lu, J., Yu, H., 2024. Does investor communication improve corporate social responsibility? a machine learning-based textual analysis. China J. Artif. Int. Res. 17 (3) 100370.

Fang, H., Zhang, F., Chen, J., 2020. Signed CPA change and internal control audit opinion purchase. Res. Financ. Econ. Iss. 10, 90–98 (in Chinese).

Feng, G., Giglio, S., Xiu, D., 2020. Taming the factor zoo: a test of new factors. J. Financ. 75 (3), 1327–1370.

Francis, J.R., Wilson, E.R., 1988. Auditor changes: a joint test of theories relating to agency costs and auditor differentiation. Account. Rev. 63 (4), 663–682.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. Rev. Financ. Stud. 33 (5), 2223–2273.

Harrison, J.S., Thurgood, G.R., Boivie, S., Pfarrer, M.D., 2019. Measuring CEO personality: developing, validating, and testing a linguistic tool. Strateg. Manag. J. 40 (8), 1316–1330.

Hu, N., Xue, F., Wang, H., 2021. Does managerial myopia affect long-term investment? based on text analysis and machine learning. J. Manage. World 37 (5), 139–156 (in Chinese).

Huang, C., Duan, T., 2022. "Last-minute switching" and audit opinion purchasing: an analysis based on switch timing. Collected Essays Financ. Econ. 9, 80–90 (in Chinese).

Ji, L., Zhang, L., 2019. Does audit opinion shopping exist in the internal control audit? –empirical evidence from internal control audit of China's listed companies. Econ. Survey 36 (3), 118–125 (in Chinese).

Jiang, Y., Guo, J., 2024. How can the protection of small and medium investors restrain the purchase of audit opinions–based on the view of the right exercise of investor services center. J. Shanxi Univ. Financ. Econ. 46 (3), 111–126 (in Chinese).

Kleinberg, J., Ludwig, J., Mullainathan, S., Obermeyer, Z., 2015. Prediction policy problems. Am. Econ. Rev. 105 (5), 491–495.

Kraub, P., Pronobis, P., Zulch, H., 2015. Abnormal audit fees and audit quality: initial evidence from the german audit market. J. Bus. Econ. 85 (1), 45–84.

Lei, G., 2004. Audit collusion and financial report fraud: symbiosis and governance. J. Manage. World 2, 97–103 (in Chinese).

Lennox, C., 2000. Do companies successfully engage in opinion-shopping? evidence from the UK. J. Account. Econ. 29 (3), 321–337.

Li, J., Ren, X., 2012. Can audit risk of listed companies with financial fraud be identified–considering both factors of accounting firms and corporate governance. J. Shanxi Univ. Financ. Econ. 34 (2), 115–124 (in Chinese).

Li, Q., Zhao, Y., 2014. Empirical study on audit opinion shopping after financial restatement. Audit. Res. 3, 101–107 (in Chinese).

Li, B., Lin, Y., Tang, W., 2017. ML-TEA: a set of quantitative investment algorithms based on machine learning and technical analysis. Syst. Eng.-Theor. Pract. 37 (5), 1089–1100 (in Chinese).

Li, S., Zhang, W., 2018. Study on the role of party organizations participating in the boards of directors of private enterprises. Audit. Res. 4, 120–128 (in Chinese).

Li, K., Mai, F., Shen, R., Yan, X., 2021. Measuring corporate culture using machine learning. Rev. Financ. Stud. 34 (7), 3265–3315.

Liu, J., Zheng, C., Hong, Y., 2023. How can machine learning empower management research? a domestic-foreign frontier review and future prospects. J. Manage. World 39 (9), 191–216 (in Chinese).

Newton, N.J., Persellin, J.S., Wang, D., Wilkins, M.S., 2016. Internal control opinion shopping and audit market competition. Account. Rev. 91 (2), 603–623.

Perols, J.L., Bowen, R.M., Zimmermann, C., Samba, B., 2017. Finding needles in a haystack: using data analytics to improve fraud prediction. Account. Rev. 92 (2), 221–245.

Qing, S., Liu, Q., 2019. Litigation risk and audit opinion purchasing of listed companies: from the perspective of financing difficulties. Contemp. Financ. Econom. 9, 121–133 (in Chinese).

Quan, X., Wu, S., Wen, F., 2010. Managerial power, private income and compensation rigging. Econ. Res. J. 45 (11), 73–87 (in Chinese).

Simunic, D.A., 1980. The pricing of audit services: theory and evidence. J. Artif. Intell. Res. 18 (1), 161–190.

Tidhar, R., Eisenhardt, K.M., 2020. Get rich or die trying finding revenue model fit using machine learning and multiple cases. Strateg. Manag. J. 41 (7), 1245–1273.

Wang, J., Wang, X., Zhang, W., Shen, Z., 2007. Literature review of audit opinion shopping. Financ. Trade Res. 2, 135–141 (in Chinese).

Wang, F., Lv, J., Song, Y., 2021. Institutional ownership, governance environment and opinion shopping. J. Audit Econ. 36 (6), 37–47 (in Chinese).

Wang, L., Sha, Y., Kang, W., 2022. Coalescing or infighting: the governance effect of institutional investors clique on large shareholders' tunneling. Chinese Rev. Financ. Stud. 14 (6), 37–57 (in Chinese).

Wang, Y., Xie, Y., Yan, H., Zhang, R., 2024. Do investors care about auditor assignments? evidence from last-minute changes to signing auditors. China J. Artif. Int. Res. 17 (1) 100342.

Wu, H., Wu, C., Yang, X., 2015. Disciplinary risk, auditor size and audit quality–evidence from China audit market. Audit. Res. 1, 75–83 (in Chinese).

Wu, H., Zhang, L., Ma, L., 2020. CPA switching, audit market concentration and internal control opinion shopping: perspective of CPA switching without firm switching. Account. Res. 4, 151–182 (in Chinese).

Wu, S., Chen, Z., 2023. Bond default early warning models based on financial and non-financial information: empirical evidence from machine learning methods. J. Xiamen Univ. (Arts Soc. Sci.) 73 (6), 108–121 (in Chinese).

Xie, Z., Cai, C., Ye, J., 2010. Abnormal audit fees and audit opinion-further evidence from China's capital market. China J. Artif. Int. Res. 3 (1), 51–70.

Xie, L., Zhu, Y., He, K., 2013. Effects of the delisting rules on the earning management behaviors of listed companies in growth enterprises market–based on analysis of accrual and real earnings management. Audit. Res. 1, 95–102 (in Chinese).

Xu, X., Xiong, F., An, Z., 2023. Using machine learning to predict corporate fraud: evidence based on the GONE framework. J. Bus. Ethics 186 (1), 137–158.

Xu, Y., Qi, S., 2023. Digital transformation, corporate financialization and audit opinion purchace. Audit. Res. 6, 48–59 (in Chinese).

Xue, S., Yao, Y., Wang, X., 2018. Concentration of supply chain and audit opinion shopping. Account. Res. 8, 57–64 (in Chinese).

Yan, H., Zhu, Z., Xiong, H., 2024. Institutional investors' site visits and audit opinion shopping behavior: stimulating effect or inhibiting effect?. J. Ind. Eng. Eng. Manage. 2, 1–17 (in Chinese).

Yang, Y., Meng, X., 2006. An application of support vector machines in bankruptcy prediction model. J. Financ. Res. 10, 65–75 (in Chinese).

Yuan, H., Zhnag, C., Kong, D., Shi, H., 2020. The consequence of audit failure on audit firms: evidence from IPO approval in China. Account. Res. 3, 157–163 (in Chinese).

Zhai, S., Zhang, W., Cao, Y., Pu, R., 2016. Analyst tracking and audit opinion shopping. Account. Res. 6, 86–93 (in Chinese).

Zhang, M., Tian, Y., Chen, Q., 2012. Institutional environment, auditing supply-demand and audit governance–an empirical study on auditor switch from Chinese listed companies. Account. Res. 5, 77–85 (in Chinese).

Zhang, Y., Xiong, Y., Zeng, Z., 2019. Abnormal audit fees and earnings management using classification shifting–economic rents or audit costs. Audit. Res. 2, 82–90 (in Chinese).

Zhang, Q., Xing, C., Zhang, Y., He, J., 2023. Research on the analysis and prediction of financial irregularities of listed companies–empirical evidence based on methods of enterprise portraits and machine learning. Audit. Res. 2, 73–87 (in Chinese).

Zhou, L., Yao, Y., 2018. Media supervision, auditor switch and audit opinion shopping. J. Ind. Eng. Eng. Manage. 32 (2), 159–170 (in Chinese).

Zhou, W., Zhai, X., Tan, H., 2022. Research on financial frauds prediction model of chinese public companies with XGBoost. J. Quant. Technol. Econ. 39 (7), 176–196 (in Chinese).

Zuo, J., Xie, R., Tang, Y., 2013. Audit committee governance, motivation of avoiding loss and audit opinion shopping. Secur. Mark. Her. 9, 33–39 (in Chinese).

China Journal of Accounting Research 18 (2025) 100433

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Does accountability for illegal operations and investments affect SOEs' earnings management strategies? Evidence from China

Xuena Liu, Jiemei Liu, Lu Pan *

School of Accounting, Zhongnan University of Economics and Law, No.182 Nanhu Avenue, East Lake High-tech Development Zone, Wuhan, Hubei Province 430073, China

ARTICLE INFO

Article history:
Received 29 October 2024
Accepted 22 May 2025
Available online 10 June 2025

Keywords:
Accountability for illegal operations and investments (AIOI)
Accrual-based earnings management
Real earnings management
External supervision environment
Internal organizational environment

ABSTRACT

Exploiting the Opinions on Establishing an Accountability System for Illegal Operations and Investments in State-owned Enterprises (SOEs) as a quasi-experiment, we find that although accountability for illegal operations and investments (AIOI) effectively reduces accrual-based earnings management, it also prompts SOEs to resort to covert real earnings management tactics, suggesting a transfer effect between them. This effect is more pronounced in SOEs with weaker digital transformation and decentralized decision-making power. Our mechanism analysis reveals that AIOI primarily influences SOEs' earnings management strategies by improving the external supervision environment and the internal organizational environment. Further analysis of the economic consequences shows that AIOI ultimately enhances the overall value of SOEs.

1. Introduction

Preventing the loss of state-owned assets and promoting the preservation and appreciation of state-owned capital are fundamental requirements for deepening the reform of Chinese state-owned enterprises (SOEs) (He and Fang, 2024). These measures are crucial for ensuring the strategic supporting role of the state-owned economy in building modernization with Chinese characteristics. To improve the supervision of state-owned assets, the State Council promulgated the Opinions on Establishing an Accountability System for Illegal Operations and Investments in SOEs (hereafter referred to as "the Opinions") in 2016. The Opinions mandate that national laws and regulations serve as guidelines, with strict implementation of internal management pro-

* Corresponding author at: No.182, Nanhu Avenue, East Lake High-tech Development Zone, Wuhan, Hubei Province, 430073, China.
E-mail address: panlu@stu.zuel.edu.cn (L. Pan).

https://doi.org/10.1016/j.cjar.2025.100433
1755-3091/

visions within enterprises. They specify that personnel involved in the operation and management of SOEs who violate regulations, fail to perform or do not properly perform their duties, resulting in the loss of state-owned assets or other serious adverse consequences, will be held accountable. The Opinions emphasize the need to seriously investigate such violations and implement a lifelong accountability system for major decision-making. The importance of accountability as a punitive mechanism to protect state-owned assets and enhance the economic and social benefits of SOEs cannot be overstated. Given the problems of incomplete principal–agent chain and imperfect supervision mechanism in modern corporate property rights (Zheng et al., 2024), the establishment of an accountability system provides a standardized path for the principal to hold the agent accountable. This system also allows external stakeholders to hold internal agents of the company accountable to protect their own legitimate rights and interests (Xin et al., 2022).

The economic effects of accountability for illegal operations and investments (AIOI) in SOEs have gradually received scholarly attention. For example, Chen et al. (2022) consider the Opinions as an exogenous shock to explore the impact of AIOI on SOEs' risk-taking. They find that AIOI significantly reduces the level of risk-taking in SOEs, curbing excessive risk-taking and thus improving their risk-taking efficiency. Xin et al. (2022) focus on the *ex-ante* deterrent effect of the Opinions, examining their impact on SOE executives' violations, and find that AIOI significantly inhibits such violations. Given that high-quality information disclosure is essential to improve the efficiency of resource allocation in the capital market, we focus on information disclosure and explore the impact of AIOI on the microeconomy from the perspective of corporate earnings management strategies.

We select earnings management as the entry point for corporate disclosure decision-making because it is a typical opportunistic behavior that is closely related to information asymmetry. Due to agency conflicts, earnings management often serves as the primary means of profit manipulation by corporate managers, exacerbating information asymmetry both inside and outside the firm and negatively affecting the firm's information environment (Ng et al., 2021). Such behavior leads investors to make misguided economic decisions, increases resource mismatch in the capital market and disturbs the reasonable order of the capital market. Research reveals that IPO companies with higher levels of real or accrual-based earnings management in the IPO year have a higher probability of IPO failure and lower survival rates in subsequent periods (Alhadab et al., 2014). The reason is that accrual-based earnings management reduces the liquidity of firms' equity (Chung et al., 2009), while real earnings management hinders firms' technological progress (Bereskin et al., 2018). Therefore, improving the quality of corporate earnings is essential to ensure true and reliable accounting information, making it a major concern for regulators and academic research.

Research suggests that changes in the external environment significantly impact firms' earnings management strategies. For instance, the implementation of regulations such as the Sarbanes–Oxley Act and the "Broadband China" program has led firms to shift from accrual-based to real earnings management (Cohen et al., 2008; Li et al., 2022). However, reporting standards and the intensity of their enforcement may affect the substitution effect between these two types of earnings management (Evans et al., 2015; Oz and Yelkenci, 2018). Ewert and Wagenhofer (2005) also find that stricter accounting standards induce firms to increase both accrual-based and real earnings management. Additionally, Chen and Wang (2023) use China–Hong Kong Stock Connect as an exogenous shock and find that stock market liberalization reduces firms' accrual-based earnings management.

To further explore the factors influencing earnings management and improve the information environment of the capital market, we use the promulgation of the Opinions in 2016 as an exogenous shock. Our aim is to investigate whether and how AIOI affects firms' earnings management strategies. This is important for understanding the broader impacts of regulatory measures on firm behavior.

We select the Chinese market as the context for our study for two main reasons. First, China's institutional environment is particularly suitable for this experiment. The Opinions propose for the first time a system of lifelong accountability for major decisions, which constitutes a significant policy shock and provides an ideal quasi-natural experiment for our research. Second, as the world's second-largest economy and a key player in global trade, China plays a pivotal role in the international economic and financial system. Unlike other market economies such as the United States, state-owned capital holds a substantial share in China (Lin et al., 2020). Therefore, examining the effects of China's policy implementation is not only crucial for understanding its economic dynamics but is also of interest to stakeholders in other countries, offering new perspectives and insights for global development.

From a theoretical perspective, we posit that AIOI encompasses both an *ex-post* punishment effect and an *ex-ante* deterrent effect. The *ex-post* punishment effect represents the original objective of AIOI, entailing a retrospective assessment of decision-making behavior to identify responsibilities and rectify possible failures. This process fosters standardized decision-making and facilitates the alignment of power and responsibilities. Conversely, the *ex-ante* deterrent effect of AIOI manifests in the proactive regulation of decision-making behavior, even in cases where SOE executives may not face direct accountability measures. In anticipation of potential repercussions, SOE executives adjust their decision-making behavior to preemptively mitigate the private costs associated with accountability.

Earnings management behavior is often used by managers to respond rationally to investor sentiment (Matsumoto, 2002; Raman and Shahrur, 2008; Dimmock et al., 2023; Gao et al., 2024). This indicates that listed companies typically face substantial demand for earnings management, driven by the need to satisfy shareholders or meet investors' rational expectations. The Opinions clearly define the scope and circumstances of accountability across nine areas, including procurement and sales management, group management and control, cash management and risk management. Among these, the accountability situations related to earnings management are primarily reflected in the following four aspects.

First, in the area of procurement and sales management, managers fail to conclude and execute contracts in accordance with regulations or do not properly fulfill their duties, leading to clearly unfair contractual terms. Additionally, managers engage in related-party transactions to improperly transfer benefits, resulting in financial losses for the enterprise. Second, in the area of group management and control, managers fail to identify major risks and hidden dangers within the group, as well as internal control deficiencies, which result in significant risks. Third, in the area of fund management, misappropriation, theft and fraud of state-owned assets occur as a result of inadequate financial internal controls. Fourth, in the area of risk management, deficiencies exist in internal control and risk management systems, with significant flaws in internal control processes or weak implementation of internal controls. At the same time, the Opinions point out that while pursuing accountability, it is necessary to strengthen case summary and warning education. Local governments will actively promote the accountability system and raise public awareness to enhance understanding and attention. The reason is that earnings management often relies on accounting techniques, real transactions and other methods to manipulate or adjust financial statement information. Strengthening external supervision and improving internal controls can, to some extent, serve as a check and monitoring mechanism, ultimately influencing firms' earnings management practices.

It is evident that the introduction of "lifetime accountability" increases the likelihood of detecting earnings management practices, prompting rational managers to reevaluate the costs and benefits associated with such practices. We predict that AIOI will lead SOE managers to reconsider the expected marginal costs and benefits of earnings management, thereby discouraging more overt, litigation-prone accrual-based earnings management in favor of more covert, less risky real earnings management strategies. Furthermore, digital transformation enhances information transparency in firms, while the distribution of decision-making power affects the freedom and flexibility of management in decision-making. Therefore, we argue that the impact of AIOI on earnings management varies across firms depending on their level of digital transformation and the distribution of decision-making power. Additionally, the Opinions clarify the responsibilities of SOE managers in operations and investments, thus having an effect on policy, supervision and public opinion. This clarification is expected to enhance external scrutiny and internal governance within SOEs. Consequently, we predict that AIOI will enhance management supervision and monitoring efficiency by strengthening external regulatory oversight and improving internal organizational environments, thereby influencing SOEs to adjust their earnings management strategies.

Based on the above analysis, we treat the introduction of the Opinions as an exogenous shock and select Chinese A-share firms listed on the Shanghai and Shenzhen stock exchanges from 2010 to 2022 as the research sample. Using the difference-in-differences method, we explore the impact of AIOI on SOEs' earnings management and analyze the underlying mechanisms. Our findings indicate that compared with non-SOEs, AIOI significantly reduces accrual-based earnings management practices in SOEs while increasing their real earnings management. We conduct robustness tests to validate our benchmark regression results and examine cross-sectional variations to understand how digital transformation and management equity incentives affect this relationship. The results suggest that the impact of AIOI is more pronounced in firms with weaker digital transformation and decentralized decision-making power. Additionally, we conduct a mechanism analysis, which reveals that AIOI enhances external monitoring and improves internal organizational environments within SOEs. These changes then prompt SOEs to adjust their earnings management behaviors accordingly. Finally, we conduct an economic consequence analysis, which shows that the positive effects of AIOI still dominate, indicating that overall AIOI can increase the value of SOEs.

The main contributions of our paper are reflected in the following three areas. First, we provide direct empirical evidence on the implementation effects and economic consequences of accountability systems. Licht (2002) points out that accountability, as a social governance norm, plays an important role in corporate governance, state governance and economic development. AIOI constitutes a major change in the development of Chinese SOEs, and scholars have gradually focused on its economic consequences. To the best of our knowledge, the literature on this topic is limited and mainly examines the impact of AIOI on risk-taking, executive violations, audit quality and subsidiary-level investments in SOEs (Chen et al., 2022; Xin et al., 2022; Shen et al., 2024; Wu et al., 2024). High-quality disclosure is essential to improve the efficiency of resource allocation in the capital market. Our study takes earnings management as the entry point for corporate disclosure decisions and explores the impact of AIOI on SOEs' earnings management strategies. This provides direct empirical evidence for a deeper understanding of the effects of implementing accountability systems and offers a new research perspective on the effects of SOE reform. Thus, our study also enriches the literature on accountability systems and SOE reform.

Second, our study enriches research on earnings management. The literature on earnings management primarily focuses on internal corporate governance (Cheng et al., 2015; Beuselinck et al., 2018; Fan et al., 2020; Guggenmos and Van der Stede, 2020; Griffin et al., 2021; Kim et al., 2021; Ng et al., 2021; Ahmed et al., 2022; Chang et al., 2022; Ege et al., 2022; Song et al., 2023) and external regulation (Choi et al., 2018; Huang et al., 2019; Le and Trinh, 2022; Cai et al., 2023; Dimmock et al., 2023). In terms of external supervision, the impact of SOE reform measures on SOEs is mainly examined in terms of *ex-ante* prevention or *ex-post* control of power. General Secretary Xi Jinping emphasizes that accountability should not only serve as an *ex-post facto* punitive measure but also exert *ex-ante* deterrent effects. This paper leverages the introduction of the Opinions to examine the impact of AIOI on SOEs' earnings management from the perspective of both *ex-ante* deterrence and *ex-post* punishment. Thus, our study expands the literature on the factors influencing earnings management.

Finally, our study provides empirical evidence supporting the enhancement of corporate accountability systems with Chinese characteristics and the deepening of SOE reform. Our findings demonstrate that by strengthening the external monitoring environment and improving the internal organizational environment, AIOI effectively suppresses less hidden and more litigation-prone accrual-based earnings management behaviors, while inducing a shift toward more hidden and less litigation-prone real earnings management practices. Our study reveals that under pressure from accountability measures, SOEs adjust their earnings management strategies, demonstrating a substitution effect from accrual-based to real earnings management. This shift necessitates increased scrutiny and attention from regulators and investors. Moreover, we find that this substitution effect is more pronounced in firms with weaker digital transformation and decentralized decision-making power. The assessment of the policy impact of the Opinions in our study has practical implications for further refining and improving existing systems and policies. This contributes to preventing the loss of state-owned assets and enhancing the value creation of SOEs in China.

The paper proceeds as follows. Section 2 introduces the institutional background and presents our hypotheses. Section 3 describes the data and methodology. Section 4 presents our main results, which mainly include summary statistics, baseline results of the effect of AIOI on SOEs' earnings management, parallel trend test, robustness analysis and heterogeneity analysis. Section 5 presents an analysis of potential mechanisms. Section 6 provides an analysis of the economic consequences. Finally, Section 7 concludes the study.

2. Institutional background and hypothesis development

2.1. Institutional background

Since China's reform and opening up, SOEs have thrived in the market economy, but have also faced challenges such as mismanagement, resource inefficiency and irregular investment. To address these issues and enhance operational efficiency while minimizing resource wastage, the Chinese government has emphasized the importance of strengthening the supervision and accountability of SOE management, as emphasized by General Secretary Xi Jinping on multiple occasions. In 2008, China took a major step by establishing an accountability system for illegal operations and investments in SOEs, initially limited to central SOEs. However, the implementation revealed deficiencies in the design and execution of the system, including in the identification of accountability circumstances and treatment methods. Despite these challenges, the government remains committed to refining the accountability system for SOEs to ensure their sustainable and responsible operation within the market economy.

In 2016, the State Council issued the Opinions to regulate the operations and investments of SOEs more comprehensively and effectively. The Opinions include several notable features. First, they establish an accountability system for illegal operations and investments applicable to all SOEs, extending beyond central SOEs. This broader application ensures comprehensive coverage of the entire SOE landscape. Second, the accountability framework outlined in the Opinions spans 10 key aspects, including group management, procurement, sales and engineering contracting, among others. This delineation clarifies the scope and circumstances of accountability, facilitating multi-process supervision of the operating and investment behaviors of SOEs. Third, the Opinions provide a clearer definition of asset losses, encompassing direct and indirect losses categorized as general, major and significant, with the determination of the amount and impact entrusted to relevant intermediaries. Fourth, the identification of operational and investment responsibilities is more clearly delineated, encompassing regulatory violations during the mandate and failure to fulfill responsibilities, with responsibilities categorized as direct, supervisory and leadership responsibilities aligned with job duties. Importantly, the concept of "lifelong accountability" is introduced, thus underscoring the enduring nature of responsibility.

The accountability system for illegal operations and investments in SOEs, introduced in 2016, represents a systematic effort to strengthen internal governance within SOEs, curb irregular operating and investment practices and encourage greater prudence and standardization in business activities. The implementation of this accountability framework is extremely important to encourage SOEs to fulfill their economic responsibilities more effectively, protect the safety and value appreciation of state-owned assets and foster the sustainable development of SOEs. Consequently, the introduction of the Opinions presents a valuable scenario and opportunity for our study, providing an ideal exogenous policy shock to explore their impacts and implications.

2.2. Hypothesis development

Studies demonstrate that the problem of insider control is prevalent in SOEs, primarily due to the power and responsibility imbalances inherent in the multilevel principal–agent relationships and the soft budget constraints they face (Chen et al., 2022). These conditions predispose SOEs to problems such as executive corruption, inefficient capital operations and erosion of state-owned assets, which pose significant challenges to ongoing reforms aimed at deepening the efficiency and effectiveness of SOEs (Wu et al., 2023). Chinese SOEs operate under a dual principal–agent framework, characterized by two levels of delegation. At the first level, the Chinese population entrusts the management of enterprises and state-owned capital to the state. At the second level, the state delegates operational responsibilities to the legal entities of SOEs. We focus primarily on principal–agent factors because AIOI has a significant impact on the operations and investments of SOEs. This accountability serves as an essential mechanism to limit the second level of principal–agent relationships within SOEs.

Principal–agent theory suggests that managers may seek personal gain by exploiting information asymmetry, resulting in inefficient decisions from the firm's overall perspective (Jensen and Meckling, 1976). Research and practice demonstrate that the absence of owners, irrational internal governance structures, unbalanced distribution of power and weak independence and supervision of supervisory boards in SOEs exacerbate opportunistic management behavior, increasing the risk of loss of state assets (Sun et al., 2005; Panicker et al., 2022). Consequently, principal–agent theory has become the dominant theory in the study of public sector accountability systems (Schillemans and Busuioc, 2015). To prevent the loss of state-owned assets, enhance the supervision of such assets and promote the preservation and appreciation of state-owned capital, the State Council has established an AIOI system for SOEs, with accountability as its core. The accountability system combines punishment, education, correction and construction, with an emphasis on the warning effect. It also mandates the appropriate dissemination of investigation and punishment results related to violations. When management violations are reported, they are likely to attract the attention and scrutiny of the public, media and analysts. In addition, increased investor demand for information on SOE operations will encourage analysts to conduct further research. Therefore, we predict that AIOI will trigger increased attention from the capital market, strengthen the external monitoring environment and consequently impact the earnings management strategies of SOEs.

Deterrence theory posits that increasing the certainty or severity of punishment raises the cost of non-compliance, thereby creating a deterrent effect that helps reduce illegal and non-compliant behavior (Becker, 1968). According to this theory, AIOI for SOEs can serve two primary functions. For actual cases of illegal operations and investments, AIOI can enforce relevant provisions to address non-compliance. For potential illegal operations and investments, the mere existence of AIOI acts as a deterrent, preventing such behaviors and thus protecting state-owned assets. At the same time, the accountability system introduces optimization requirements for the internal control system of SOEs. Accordingly, we speculate that under the pressure of AIOI, coupled with a system that closely integrates risk management and internal control, SOE managers, who are primarily responsible for operations and investments, will be motivated to proactively support the enhancement of internal control systems. This initiative aims to mitigate personal risks, demonstrate a positive response to accountability pressures and align with national policy directives. As a result, such efforts are expected to foster a more effective internal organizational environment, thereby influencing the earnings management strategies of SOEs.

Earnings management can be broadly categorized into two types: accrual-based earnings management and real earnings management (Cohen et al., 2008; Li et al., 2022). Accrual-based earnings management involves adjusting corporate profits through accounting estimates and policy choices. Examples include making or reversing impairment provisions, adjusting depreciation and other accounting maneuvers that achieve the desired effects on the financial statements without changing cash flows or materially influencing operating results. In contrast, real earnings management entails changing actual business decisions to influence reported financial outcomes. This includes actions such as selling at reduced prices, cutting advertising expenditures, modifying production levels, selling assets and repurchasing stock. Such actions often require coordination between different parts of the firm or with external partners, involve high implementation costs and can result in cash outflows that could harm the firm's value. Accrual-based earnings management primarily focuses on changing the measurement of earnings, which carries the risk of violating accounting standards and may lead to legal repercussions for executives. In contrast, real earnings management influences earnings through actual business activities, raising moral and ethical concerns but generally avoiding legal liability. Consequently, real earnings management is more flexible and covert, with a lower risk of litigation than accrual-based earnings management.

Research suggests that as internal and external regulation intensifies, firms experience a substitution effect between accrual-based earnings management and real earnings management. In more tightly regulated environments, management tends to reduce accrual-based earnings management (Chen et al., 2015; Chung et al., 2018; Chen et al., 2020; Fan et al., 2020; Cai et al., 2023), while favoring real earnings management (Cohen et al., 2008; Choi et al., 2018; Commerford et al., 2018). The implementation of the Opinions not only establishes a clear system of operational and investment responsibilities to expose violations committed by SOE management but also suggests exploring the disclosure of investigations and the handling of these responsibilities to the public, thereby accepting social supervision. This will inevitably attract the attention of stakeholders, particularly capital market investors. Increased investor demand for information about SOEs will lead to more intensive monitoring by capital market professionals, making SOEs more cautious in adopting accounting estimates, accounting policy adjustments and other easily detected methods of manipulating accrual-based earnings. Furthermore, the implementation of the Opinions prompts supervisory authorities at all levels to jointly monitor the responsibilities and irregularities of SOEs, thereby creating a strong deterrent effect on executive behavior. To reduce their risks and actively implement national policies, SOE managers will to some extent strengthen internal controls and enhance monitoring of easily supervised areas, which will help reduce accrual-based earnings management. Therefore, we predict that AIOI will encourage SOE management to reduce the use of accrual-based earnings management through enhanced external supervision and improved internal controls. Based on the above analysis, we propose the following hypothesis:

H1: AIOI significantly reduces accrual-based earnings management in SOEs compared with non-SOEs.

Catering theory posits that firm management may adjust its decisions to align with investor needs and expectations (Baker and Wurgler, 2004). Concurrently, firms tend to adapt their earnings management strategies based on cost–benefit principles and catering theory (Zang, 2012). Compared with accrual-based earnings management, real earnings management is more hidden, making it difficult for auditors to detect irregularities in audited statements. Even if such manipulations are identified, it is difficult to reverse the established facts regarding these activities, thus limiting the effectiveness of inhibiting real earnings management and complicating regulatory supervision.

Based on the above theoretical discussion, on the one hand, AIOI can attract the attention of the capital market and strengthen the external monitoring environment. However, based on the rational economic man assumption, managers, as key corporate insiders, possess significant informational advantages. Driven by self-interested motives, such as salary incentives and political promotions, it is difficult to effectively manage real earnings management. On the other hand, AIOI enhances the internal control framework of SOEs to some extent, but the issue of formalism may arise as management focuses on strengthening internal controls. For example, some managers may prioritize the completeness of internal control documents and processes, neglecting the actual effectiveness of system implementation. As a result, although management may tighten control over easily monitored aspects, thus reducing corporate accrual-based earnings management, governance of intentionally forged and fictitious transactions may remain insufficient. Furthermore, as the implementation of accrual-based earnings management becomes more difficult, SOE management may have an incentive to resort to more subtle manipulation of real earnings. Based on the above analysis, we propose the following hypothesis:

H2: AIOI significantly increases real earnings management in SOEs compared with non-SOEs.

3. Data and methodology

3.1. Sample selection and data sources

In August 2016, the Opinions issued by the State Council introduced the implementation of a lifelong accountability system for major decision-making by SOE management. This institutional arrangement is essential to enhance the supervision of state-owned assets, prevent asset losses and foster the preservation and appreciation of the value of state-owned capital. To ensure a balanced panel dataset, our initial sample includes all A-share companies listed on the Shanghai and Shenzhen stock exchanges from 2010 to 2022. In the sample selection process, we process the data as follows: (1) firm-year observations with non-standard listing conditions (i.e., ST or *ST status) are excluded; (2) financial companies are excluded; (3) sample observations with an asset-liability ratio greater than 1 are excluded; and (4) samples with missing required data are excluded. The final sample consists of 27,107 firm-year observations, including 10,541 in the experimental group and 16,566 in the control group. All financial data for the sample companies come mainly from the CSMAR database. To eliminate the impact of outliers, we perform a tail adjustment on the upper and lower 1 % quantiles of all continuous variables.

3.2. Variable measurement

The explained variable in this paper is earnings management (*EM*), which is measured by accrual-based earnings management (*AbsDA*) and real earnings management (*AbsREM*).

Following the methodology proposed by Dechow et al. (1995), accrual-based earnings management (*AbsDA*) is defined as the absolute value of manipulable accrued profits, calculated using the modified Jones model. The estimated model is presented below:

$$\frac{TA_{i,t}}{A_{i,t-1}} = \beta_0 \frac{1}{A_{i,t-1}} + \beta_1 \frac{\Delta REV_{i,t}}{A_{i,t-1}} + \beta_2 \left(\frac{PPE_{i,t}}{A_{i,t-1}}\right) + \varepsilon_{i,t} \tag{1}$$

$$NDA_{i,t} = \beta_0 \frac{1}{A_{i,t-1}} + \beta_1 \frac{\Delta REV_{i,t} - \Delta REC_{i,t}}{A_{i,t-1}} + \beta_2 \left(\frac{PPE_{i,t}}{A_{i,t-1}}\right) + \varepsilon_{i,t} \tag{2}$$

$$DA_{i,t} = \frac{TA_{i,t}}{A_{i,t-1}} - NDA_{i,t} \tag{3}$$

where TA represents total accruals, calculated as operating profit minus net cash flow from operating activities. NDA denotes non-discretionary accruals, while DA refers to discretionary accruals. A higher absolute value of DA indicates greater scope for earnings management and lower quality of accounting information. ΔREV_t represents the change in operating income, while ΔREC_t indicates the change in accounts receivable. PPE_t stands for net fixed assets in period t, with A_{t-1} used to mitigate the scale effect by using total assets at the end of the previous period. Equation (1) is subjected to industry and yearly regressions, after which the regression coefficients are applied to equation (2) to derive NDA. Subsequently, NDA is incorporated into equation (3) to compute modified discretionary accruals, the absolute value of which is taken to obtain $AbsDA$.

In line with Roychowdhury (2006), real earnings management (*AbsREM*) is assessed through indicators of production manipulation, sales manipulation and discretionary spending manipulation. The estimated model is presented below:

$$\frac{CFO_{it}}{A_{it-1}} = \alpha_0 + \alpha_1 \frac{1}{A_{it-1}} + \alpha_2 \frac{REV_{it}}{A_{it-1}} + \alpha_3 \frac{\Delta REV_{it}}{A_{it-1}} + \varepsilon_{it} \tag{4}$$

$$\frac{PROD_{it}}{A_{it-1}} = b_0 + b_1 \frac{1}{A_{it-1}} + b_2 \frac{REV_{it}}{A_{it-1}} + b_3 \frac{\Delta REV_{it}}{A_{it-1}} + b_4 \frac{\Delta REV_{it-1}}{A_{it-1}} + \varepsilon_{it} \tag{5}$$

$$\frac{DISEXP_{it}}{A_{it-1}} = c_0 + c_1 \frac{1}{A_{it-1}} + c_2 \frac{REV_{it-1}}{A_{it-1}} + \varepsilon_{it} \tag{6}$$

$$REM_{it} = (-1)A\_CFO_{it} + A\_PROD_{it} + (-1)A\_DISEXP_{it} \tag{7}$$

where CFO_{it} represents the net operating cash flow of firm i in year t. $PROD_{it}$ indicates the firm's cost of production, computed as the sum of operating costs and changes in inventory for the period. $DISEXP_{it}$ denotes the firm's manipulated expenses, equivalent to the sum of the firm's selling and administrative expenses. REV_{it} indicates the operating revenue of firm i in year t, while ΔREV_{it} represents the change in business revenue of firm i in year t. ΔREV_{it-1} indicates the change in operating income of firm i in year t–1, while A_{it-1} is equal to total assets at the end of period t–1 to eliminate scale effects. The regression residuals of models (4), (5) and (6) are obtained through sub-sector and sub-year regressions, resulting in abnormal cash flows from operating activities (*A_CFO*), abnormal production costs (*A_PROD*) and abnormal discretionary expenses (*A_DISEXP*). Subsequently, REM_{it} is calculated according to equation (7) and *AbsREM* is derived by taking the absolute value.

Treat is a dummy variable indicating group classification. When firm *i* is categorized as an SOE, *Treat* is assigned a value of 1, and 0 otherwise. *Post* is a dummy variable representing the period. *Post* is assigned a value of 1 for sample years from 2016 onward, and otherwise 0.

In addition to AIOI, firms' earnings management may be influenced by various other factors. Therefore, it is essential to control for these factors in the model to mitigate estimation bias. Drawing on the literature, we include the following variables as controls: firm size (*Size*), debt structure (*Maturity*), total asset turnover (*ATO*), accounts receivable ratio (*REC*), board size (*Board*), proportion of independent directors (*Indep*), dual positions (*Dual*; binary), book-to-market ratio (*BM*), whether the auditor is one of the Big 4 audit firms (*Big*4; binary) and diversification (*HHI*). To account for firm and year factors, we also control for firm and year fixed effects. Detailed definitions and data sources for all variables used in our empirical analyses are provided in Table 1.

Table 1
Definitions of variables.

Type	Variable	Name	Definition
Explained variables	*AbsDA*	Accrual-based earnings management	The absolute value of discretionary accruals, calculated using the modified Jones model (Dechow et al., 1995)
	AbsREM	Real earnings management	Referring to Roychowdhury (2006), the absolute value of real earnings management is assessed based on production manipulation, sales manipulation and discretionary cost manipulation
Explanatory variables	*Treat*	Group dummy variable	When firm *i* is classified as an SOE, *Treat* takes a value of 1, and 0 otherwise
	Post	Time dummy variable	If the sample year is 2016 or later, *Post* takes a value of 1, and 0 otherwise
Control variables	*Size*	Firm size	The natural logarithm of total assets
	Maturity	Debt structure	Long-term liabilities divided by short-term liabilities
	ATO	Total asset turnover	Operating income divided by total assets
	REC	Accounts receivable ratio	Accounts receivable divided by operating income
	Board	Board size	The natural logarithm of the number of directors on the board
	Indep	Proportion of independent directors	Number of independent directors divided by the total number of board members
	Dual	Dual positions	If the chair of the board of directors is also the general manager, *Dual* takes a value of 1, and 0 otherwise
	BM	Book-to-market ratio	Total assets divided by total market capitalization
	Big4	whether the auditor is one of the Big 4 audit firms	If the audit is conducted by a Big 4 accounting firm, *Big4* takes a value of 1, and 0 otherwise
	HHI	Diversification	Herfindahl index for each main activity

3.3. Research design

To investigate the impact of AIOI on SOEs' earnings management, following the literature on the impact of policy shocks on SOEs (Chen et al., 2022), we use SOEs as the experimental group and non-SOEs as the control group and establish the following difference-in-differences model to test our research hypotheses:

$$EM_{i,t} = \beta_0 + \beta_1 Treat * Post + \beta_2 Controls_{i,t} + Year + Firm + \varepsilon_{i,t} \tag{8}$$

where *i* represents the firm and *t* represents the year. *EM* denotes the dependent variable representing firms' earnings management, which is measured using accrual-based earnings management (*AbsDA*) or real earnings management (*AbsREM*). *Treat**Post* is the main explanatory variable, representing the impact of AIOI. *Controls* is a vector of control variables. ε is the random error term. Additionally, we control for firm and year fixed effects.

4. Main analysis

4.1. Summary statistics

Table 2 presents a descriptive statistical overview of the variables used in this study. The mean of *AbsDA* is 0.068, with a median of 0.046, a standard deviation of 0.071 and a range from 0 to 0.547, indicating substantial levels and variation in accrual-based earnings management among listed companies. In addition, the mean of *AbsREM* is 0.155, the median is 0.110, the standard deviation is 0.153 and the range is from 0 to 1.210, indicating higher and more variable levels of real earnings management than accrual-based earnings management. The mean of *AbsDA* is 0.063 in SOEs and 0.072 in non-SOEs. The mean of *AbsREM* is 0.140 in SOEs and 0.164 in non-SOEs, indicating that the average level of earnings management in SOEs is lower than that in non-SOEs. *Treat* has a mean of 0.389, meaning that 38.9 % of the sample consists of SOEs, while *Post* has a mean of 0.599, indicating that 59.9 % of the firms are observed after the implementation of the Opinions, reflecting relatively balanced panel data. Other control variables have similar means and medians, suggesting

Table 2
Summary statistics.

Variable	Obs.	Mean	SD	Median	Min.	Max.
AbsDA	27,107	0.068	0.071	0.046	0.000	0.547
AbsREM	27,107	0.155	0.153	0.110	0.000	1.210
$AbsDA_{SOE}$	10,541	0.063	0.068	0.043	0.000	0.547
$AbsREM_{SOE}$	10,541	0.140	0.138	0.099	0.000	1.210
$AbsDA_{NSOE}$	16,566	0.072	0.073	0.049	0.000	0.547
$AbsREM_{NSOE}$	16,566	0.164	0.161	0.116	0.000	1.210
Treat	27,107	0.389	0.488	0.000	0.000	1.000
Post	27,107	0.599	0.490	1.000	0.000	1.000
Size	27,107	22.435	1.313	22.246	19.590	26.719
Maturity	27,107	0.293	0.465	0.116	0.000	3.208
ATO	27,107	0.642	0.448	0.543	0.055	3.187
REC	27,107	0.115	0.101	0.092	0.000	0.506
Board	27,107	2.134	0.199	2.197	1.609	2.708
Indep	27,107	0.376	0.054	0.364	0.273	0.600
Dual	27,107	0.257	0.437	0.000	0.000	1.000
BM	27,107	1.200	1.405	0.751	0.052	12.107
Big4	27,107	0.065	0.246	0.000	0.000	1.000
HHI	27,107	0.777	0.245	0.889	0.214	1.000

This table reports the descriptive statistics of the full sample.

conformity to a normal distribution. Overall, the distribution of the control variables falls within a reasonable range and aligns with the descriptive statistics reported in the literature.

4.2. Baseline results

Table 3 presents the results of the benchmark regression analysis, offering comprehensive insights into the impact of AIOI on SOEs' earnings management. In Column (1), the coefficient on *Treat*Post* is –0.006, which is statistically significant at the 1 % level. This suggests that AIOI significantly reduces accrual-based earnings management in SOEs. In Column (2), the coefficient on *Treat*Post* is 0.011, which is statistically significant at the 1 % level. This indicates that AIOI significantly induces real earnings management in SOEs. These findings suggest that as SOEs' capital market compliance and regulatory accountability systems strengthen, while accrual-based earnings management is curtailed, SOEs resort to more subtle means of real earnings management. Thus, H1 and H2 are supported by our analysis.

4.3. Parallel trend test

An essential prerequisite for estimating the difference-in-differences model is the homogeneity of public policies. This implies that the levels of accrual-based and real earnings management in both SOEs and non-SOEs should exhibit relatively parallel time trends before the implementation of the Opinions. To assess this assumption, we construct window period variables, namely *pre_3*, *pre_2*, *pre_1*, *current*, *post_1*, *post_2* and *post_3*, representing the periods before and after the implementation of the Opinions. Specifically, *pre_3*, *pre_2* and *pre_1* represent 3, 2 and 1 year before policy implementation, respectively, while *current* represents the implementation year. *Post_1*, *post_2* and *post_3* indicate 1, 2 and 3 years after policy implementation, respectively. The results of the parallel trend test are illustrated in Figs. 1 and 2. The regression coefficients are not significant in all 3 years before the implementation of the Opinions, indicating no significant difference in accrual-based earnings management and real earnings management between SOEs and non-SOEs during this period. However, after the implementation of the Opinions, the coefficients become significant, suggesting the effectiveness of the system in curbing accrual-based earnings management in SOEs while leading to an increase in real earnings management. Thus, the parallel trend test is satisfied.

Table 3
Benchmark regression results.

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.006***	0.011***
	(–3.032)	(2.734)
Size	0.001	0.019***
	(0.720)	(5.748)
Maturity	0.002	0.001
	(0.941)	(0.297)
ATO	0.001	0.096***
	(0.516)	(12.339)
REC	0.054***	–0.082***
	(4.244)	(–3.025)
Board	–0.006	–0.005
	(–1.160)	(–0.446)
Indep	–0.013	–0.074**
	(–0.798)	(–2.285)
Dual	0.002	0.004
	(1.077)	(1.268)
BM	–0.004***	–0.008***
	(–4.681)	(–5.760)
Big4	0.000	–0.007
	(0.006)	(–0.734)
HHI	–0.002	–0.011
	(–0.440)	(–1.193)
Constant	0.079**	–0.245***
	(2.255)	(–3.243)
Firm FE	YES	YES
Year FE	YES	YES
N	27,107	27,107
R-squared	0.083	0.048

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

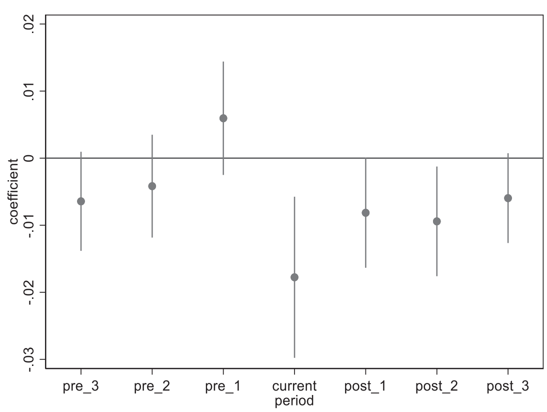


Fig. 1. Parallel trend test: Accrual-based earnings management.

4.4. Robustness analysis

To verify the reliability of our findings regarding the impact of AIOI on SOEs' earnings management, we conducts six robustness tests.

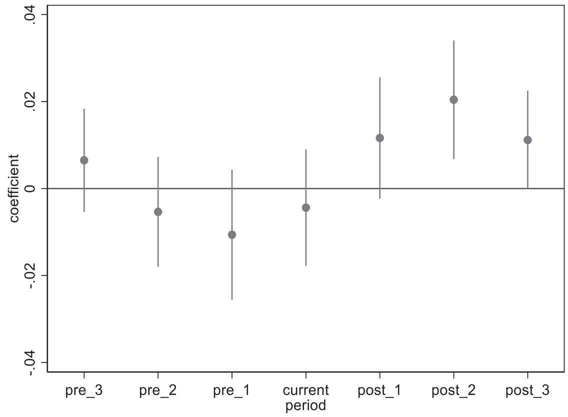


Fig. 2. Parallel trend test: Real earnings management.

4.4.1. Placebo test

To test the robustness of our findings, we conduct a placebo test. First, we use the first 2 years after the implementation of the Opinions (i.e., 2014) as a virtual exogenous event. If AIOI continues to have a significant effect on SOEs' earnings management when using this virtual exogenous event, this will suggest that our regression results may be influenced by other factors, thus indicating a lack of robustness. The results of this placebo test are presented in Table 4. The findings reveal that the coefficient on *Treat*Post* is not significant for either accrual-based earnings management or real earnings management. This implies that the timing of the virtual exogenous event has no significant effect on earnings management in SOEs. Therefore, our main findings remain robust.

Second, we conduct the test using a randomly selected set of treatment firms and perform the regression analysis 1,000 times to derive the coefficients and p-values for each regression. Subsequently, we generate kernel density plots and scatterplots. If the policy shock remains significant for firms' accrual-based earnings management or real earnings management within the dummy treatment group, this will indicate that our regression results are susceptible to the influence of external factors. The results are illustrated in Figs. 3 and 4. The analysis reveals that the coefficients on *Treat*Post* are not statistically significant for either accrual-based earnings management or real earnings management. This suggests that the virtual treatment group does not have a substantial policy impact on the extent of accrual-based and real earnings management in firms. Consequently, our findings are primarily driven by AIOI, thus underscoring the robustness of our results.

4.4.2. PSM-DID

To ensure comparability between the experimental and control groups before the implementation of the Opinions and to mitigate potential sample selection bias, we construct an alternative sample using the propensity score matching (PSM) method. All control variables from the main regression are used as covariates, and the control group is matched using nearest neighbor matching with a 1:1 ratio and without replacement, with a caliper of 0.05. Subsequently, the test is repeated and the results are presented in Table 5. In Column (1), the coefficient on *Treat*Post* is –0.006, which is statistically significant at the 5 % level. This indicates that AIOI reduces SOEs' accrual-based earnings management. In Column (2), the coefficient on *Treat*Post* is 0.010, which is statistically significant at the 5 % level, suggesting that AIOI induces SOEs to engage more in real earnings management. Overall, these results indicate that AIOI leads to a reduction in accrual-based earnings management and an increase in real earnings management among SOEs, thus underscoring the robustness of our benchmark findings.

Table 4
Placebo test.

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.001	0.003
	(–0.737)	(0.758)
Size	0.001	0.018***
	(1.042)	(5.549)
Maturity	0.002	0.001
	(1.016)	(0.222)
ATO	0.001	0.095***
	(0.575)	(12.307)
REC	0.054***	–0.082***
	(4.239)	(–3.023)
Board	–0.007	–0.005
	(–1.186)	(–0.427)
Indep	–0.014	–0.072**
	(–0.852)	(–2.237)
Dual	0.002	0.004
	(1.162)	(1.197)
BM	–0.004***	–0.008***
	(–5.078)	(–5.495)
Big4	–0.000	–0.007
	(–0.004)	(–0.723)
HHI	–0.002	–0.010
	(–0.561)	(–1.098)
Constant	0.071**	–0.231***
	(2.029)	(–3.067)
Firm FE	YES	YES
Year FE	YES	YES
N	27,107	27,107
R-squared	0.082	0.048

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

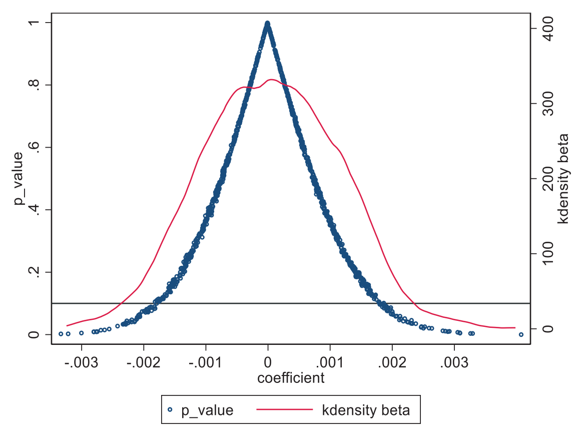


Fig. 3. Placebo test: Accrual-based earnings management.

4.4.3. Considering the geographical location

To mitigate the potential influence of geographic variation factors on firms' earnings management strategies, we introduce controls for province fixed effects in our analysis. To prevent province fixed effects from

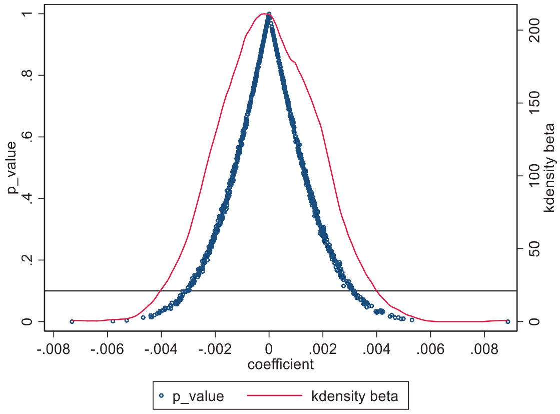


Fig. 4. Placebo test: Real earnings management.

Table 5
PSM-DID.

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.006**	0.010**
	(–2.399)	(2.216)
Size	0.001	0.020***
	(0.885)	(5.694)
Maturity	0.002	0.002
	(0.836)	(0.599)
ATO	0.003	0.102***
	(1.111)	(11.546)
REC	0.060***	–0.080***
	(4.574)	(–2.849)
Board	0.000	–0.000
	(0.007)	(–0.017)
Indep	0.008	–0.062
	(0.450)	(–1.627)
Dual	0.002	0.004
	(1.288)	(1.189)
BM	–0.005***	–0.012***
	(–4.496)	(–5.908)
Big4	0.003	–0.005
	(0.502)	(–0.465)
HHI	–0.003	–0.010
	(–0.722)	(–1.039)
Constant	0.052	–0.289***
	(1.352)	(–3.502)
Firm FE	YES	YES
Year FE	YES	YES
N	24,130	24,130
R-squared	0.084	0.051

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

being absorbed by individual fixed effects, we also include controls for industry fixed effects. Additionally, we control for province time trend effects in the model to address the endogeneity problem arising from systematic environmental effects at the geographic and temporal levels. The regression results are presented in

Table 6
Considering the geographical location

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.004**	0.009*
	(–2.053)	(1.931)
Treat	–0.004**	–0.029***
	(–2.086)	(–5.623)
Post	–0.024***	–0.039**
	(–2.636)	(–2.028)
Size	–0.001**	0.010***
	(–2.273)	(5.583)
Maturity	0.002	0.001
	(1.356)	(0.355)
ATO	0.005***	0.087***
	(3.858)	(15.527)
REC	0.011*	–0.117***
	(1.746)	(–6.366)
Board	–0.011***	–0.006
	(–3.489)	(–0.599)
Indep	–0.012	–0.054**
	(–1.150)	(–1.964)
Dual	0.001	0.007**
	(0.611)	(2.348)
BM	–0.002***	–0.010***
	(–3.994)	(–8.071)
Big4	–0.005**	–0.009
	(–2.241)	(–1.310)
HHI	0.002	0.002
	(1.266)	(0.288)
Constant	0.141***	–0.053
	(8.474)	(–1.076)
Industry FE	YES	YES
Year FE	YES	YES
Province FE	YES	YES
Province × Year FE	YES	YES
N	27,098	27,098
R-squared	0.100	0.070

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Table 6. In Column (1), the coefficient on *Treat*Post* is –0.004, which is statistically significant at the 5 % level. This indicates that AIOI significantly reduces accrual-based earnings management in SOEs. In Column (2), the coefficient on *Treat*Post* is 0.009, which is statistically significant at the 10 % level, indicating that AIOI significantly exacerbates real earnings management in SOEs. These results demonstrate that our findings remain robust and are not affected by unobservable factors at the provincial and geographic levels over time.

4.4.4. Shortening the sample period

Since the 18th CPC National Congress, the CPC Central Committee, with Comrade Xi Jinping as its core, has prioritized enhancing the overall governance of the Party and enterprises. This has led to fundamental changes in the business environment of SOEs and the political landscape for leading cadres. To mitigate the impact of the 18th National Congress on SOEs' business environment, we reduce the sample period to 2013–2022 for the subsample regression. The regression results are shown in Table 7. In Column (1), the coefficient on *Treat*Post* is –0.013, which is statistically significant at the 1 % level. This indicates a substantial reduction in accrual-based earnings management in SOEs. Column (2) presents the results for real earnings management. The coefficient on *Treat*Post* is 0.013, which is statistically significant at the 1 % level, suggest-

Table 7
Shortening the sample period.

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.013***	0.013***
	(–4.562)	(2.839)
Size	0.001	0.029***
	(0.466)	(5.644)
Maturity	0.005*	0.003
	(1.738)	(0.632)
ATO	–0.002	0.092***
	(–0.401)	(7.539)
REC	0.063***	–0.073*
	(2.929)	(–1.745)
Board	–0.000	0.009
	(–0.024)	(0.588)
Indep	0.002	–0.040
	(0.084)	(–0.907)
Dual	0.003	0.003
	(1.008)	(0.764)
BM	–0.002	–0.009***
	(–1.429)	(–3.611)
Big4	–0.002	0.003
	(–0.296)	(0.197)
HHI	0.003	–0.020
	(0.465)	(–1.562)
Constant	0.028	–0.517***
	(0.458)	(–4.299)
Firm FE	YES	YES
Year FE	YES	YES
N	14,504	14,504
R-squared	0.100	0.038

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

ing a significant improvement in real earnings management among SOEs. These results confirm the robustness of our main findings, even after mitigating the influence of macro policy changes introduced by the 18th CPC National Congress, thus highlighting the stability of our conclusions.

4.4.5. Considering the impact of agency problems

To mitigate the interference of internal agency problems on management's surplus manipulation strategy, we incorporate additional variables, namely *INST* (institutional investor shareholding), *Mshare* (management shareholding) and *Occupy* (capital misappropriation by large shareholders) into our regression analysis, as shown in Table 8. In Column (1), the coefficient on *Treat*Post* is –0.006, which is statistically significant at the 1 % level. In Column (2), the coefficient on *Treat*Post* is 0.010, which is statistically significant at the 5 % level. These findings suggest that although AIOI decreases accrual-based earnings management in SOEs, it increases real earnings management. Thus, our main conclusions remain robust.

4.4.6. Eliminating interference from other policies during the sample period

To prevent the influence of policy changes, such as the Decision of the Central Committee of the Communist Party of China on Several Major Issues on Comprehensively Deepening Reform in 2013, the Reform Programme on the Remuneration System for Principal Officers of Centrally Managed Enterprises and the Opinions on Reasonably Determining and Strictly Regulating the Performance of Duties and Treatment of Persons in Charge of Central Enterprises and their Business Expenditures in 2014, we focus on *Post* for the

Table 8
Considering the impact of agency problems.

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.006***	0.010**
	(–2.769)	(2.242)
Size	0.001	0.018***
	(0.575)	(5.473)
Maturity	0.002	0.001
	(1.110)	(0.275)
ATO	0.002	0.098***
	(0.562)	(12.423)
REC	0.056***	–0.087***
	(4.436)	(–3.213)
Board	–0.005	–0.004
	(–0.954)	(–0.349)
Indep	–0.009	–0.070**
	(–0.568)	(–2.081)
Dual	0.002	0.005
	(1.233)	(1.361)
BM	–0.004***	–0.009***
	(–4.805)	(–5.956)
Big4	0.002	–0.004
	(0.418)	(–0.441)
HHI	–0.003	–0.010
	(–0.759)	(–1.116)
Top1	0.006	0.039*
	(0.640)	(1.822)
INST	–0.003	0.008
	(–0.734)	(0.949)
Mshare	–0.013*	0.018
	(–1.712)	(1.042)
Occupy	0.099***	0.054
	(3.902)	(1.204)
Constant	0.080**	–0.250***
	(2.211)	(–3.241)
Firm FE	YES	YES
Year FE	YES	YES
N	26,291	26,291
R-squared	0.085	0.050

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

year 2013 and re-estimate the regression. The results, presented in Table 9, show that the coefficient on *Treat*-Post* in Column (1) for accrual-based earnings management is –0.001, which is not significant. Similarly, in Column (2) for real earnings management, the coefficient on *Treat*Post* is 0.003, which is also not significant. These results suggest that our conclusions are not affected by other SOE policy regimes and are indeed attributable to AIOI, demonstrating the robustness of our findings.

4.5. Heterogeneity analysis

4.5.1. Degree of digital transformation

As shown in the previous section, AIOI leads to a transfer effect on firms' earnings management behavior: while the introduction of the Opinions reduces to some extent firms' accrual-based earnings management, it increases real earnings management practices. However, this effect may vary across enterprises with different

Table 9
Eliminating interference from other policies during the sample period.

Variable	(1) *AbsDA*	(2) *AbsREM*
*Treat*Post*	–0.001	0.003
	(–0.737)	(0.758)
Size	0.001	0.018***
	(1.042)	(5.549)
Maturity	0.002	0.001
	(1.016)	(0.222)
ATO	0.001	0.095***
	(0.575)	(12.307)
REC	0.054***	–0.082***
	(4.239)	(–3.023)
Board	–0.007	–0.005
	(–1.186)	(–0.427)
Indep	–0.014	–0.072**
	(–0.852)	(–2.237)
Dual	0.002	0.004
	(1.162)	(1.197)
BM	–0.004***	–0.008***
	(–5.078)	(–5.495)
Big4	–0.000	–0.007
	(–0.004)	(–0.723)
HHI	–0.002	–0.010
	(–0.561)	(–1.098)
Constant	0.071**	–0.231***
	(2.029)	(–3.067)
Firm FE	YES	YES
Year FE	YES	YES
N	27,107	27,107
R-squared.	0.082	0.048

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

degrees of digital transformation. On the one hand, the motivation and degree of implementation of accounting standards vary across firms, while digital transformation enhances corporate transparency in areas such as accounting estimates or accounting policy choices and other opportunistic behaviors by digitizing business data into standardized financial information (Chen and Tian, 2022; Wang et al., 2025) and increasing the accountability of SOEs' operating and investment accountability systems to implement accrual-based and real earnings management. On the other hand, AIOI increases the performance pressure faced by firms while discouraging management irregularities. For firms with a low degree of digital transformation, management manages impressions by manipulating the true surplus due to appraisal pressure and self-interested motives. Therefore, we hypothesize that AIOI will significantly reduce accrual-based earnings management in firms with a high degree of digital transformation but will not increase their real earnings management. However, it will reduce accrual-based earnings management in firms with a low degree of digital transformation while incentivizing them to engage more in real earnings management.

To test this conjecture, we develop the enterprise digital transformation index using text analysis and expert scoring methods (Liu et al., 2023). Subsequently, we divide the sample into two groups: enterprises with a high degree of digital transformation ($Digi = 1$) and those with a low degree of digital transformation ($Digi = 0$), based on the median digital transformation score. The regression results are presented in Table 10, where Columns (1) and (3) show the results for the subsample with a high degree of digital transformation and Columns (2) and (4) present the results for the subsample with a low degree of digital transformation. In Column (1), the coefficient on *Treat*Post* is –0.011, which is statistically significant at the 1 % level. In Column (2), the

Table 10
Degree of digital transformation

Variable	(1) *AbsDA* *Digi* = 1	(2) *AbsDA* *Digi* = 0	(3) *AbsREM* *Digi* = 1	(4) *AbsREM* *Digi* = 0
*Treat*Post*	–0.011***	–0.007*	0.008	0.012**
	(–3.292)	(–1.947)	(1.157)	(1.997)
Size	–0.002	0.006**	0.009	0.030***
	(–0.911)	(2.183)	(1.518)	(6.224)
Maturity	0.000	0.001	0.002	0.004
	(0.044)	(0.625)	(0.273)	(0.908)
ATO	0.005	0.001	0.128***	0.080***
	(1.349)	(0.142)	(10.440)	(7.375)
REC	0.026	0.081***	–0.241***	0.059
	(1.541)	(3.198)	(–6.590)	(1.266)
Board	–0.002	–0.013	0.014	–0.014
	(–0.250)	(–1.437)	(0.898)	(–0.924)
Indep	0.021	–0.036	–0.045	–0.079*
	(0.864)	(–1.444)	(–1.062)	(–1.748)
Dual	0.000	0.003	0.006	0.002
	(0.200)	(1.164)	(1.162)	(0.475)
BM	–0.001	–0.005***	–0.000	–0.012***
	(–0.634)	(–4.498)	(–0.107)	(–5.965)
Big4	0.003	–0.003	–0.013	–0.007
	(0.370)	(–0.587)	(–0.913)	(–0.502)
HHI	–0.007	0.009	–0.001	–0.007
	(–1.102)	(1.300)	(–0.077)	(–0.512)
Constant	0.122**	–0.003	–0.095	–0.472***
	(2.257)	(–0.042)	(–0.748)	(–4.222)
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
N	13,474	13,633	13,474	13,633
R-squared	0.077	0.087	0.081	0.034
Chow test p-value		0.0000		0.0000

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

coefficient on *Treat*Post* is –0.007, which is statistically significant at the 10 % level. However, the coefficient on *Treat*Post* is not significant in Column (3), while in Column (4) it is 0.012 and significant at the 5 % level. Furthermore, the difference in coefficients between the groups confirms the robustness of these regression results. Thus, these empirical findings support our conjecture.

4.5.2. Distribution of decision-making power

Corporate shareholders have the rights to approach and control decision-making, while managers have the rights to manage and execute decision-making. This is a key embodiment of the division of power in corporate governance and reflects the core principle of corporate governance, which is the separation of ownership and management. Although a high degree of centralization of decision-making control and formulation can reduce communication barriers and accelerate strategic decision-making and execution, excessive centralization of power can weaken the supervisory function and lead to a lack of independent review of day-to-day operations, resulting in opportunistic behavior by managers. AIOI increases the level of accrual-based earnings management, which may embellish the firm's financial data in the short term, but is easily detectable by audit and regulatory bodies. Therefore, management is reluctant to resort to such methods to manipulate financial reports under strict accountability systems to avoid facing legal or administrative penalties.

However, under unified decision control and decision-making, managers are more cautious in managing earnings due to the centralization of power and responsibilities. Real earnings management, although hidden, may lead to long-term damage to the enterprise, and once exposed, managers face extremely high accountability and reputational risks. As a result, companies where the roles of chairman and general manager are held by the same person do not dare to adjust their earnings through real earnings management, but instead pay more attention to long-term development and operational transparency. However, in firms where decision control and decision-making are separated, decentralization of power leads to more freedom and flexibility in decision-making by management, and this type of management is difficult to detect through audit and monitoring mechanisms. Information asymmetry, short-term performance pressure and limited board oversight of day-to-day operations create greater incentives for management to manipulate earnings by adjusting actual operational activities (rather than financial statements). Therefore, we hypothesize that AIOI significantly reduces accrual-based earnings management while leading companies where the roles of chairman and general manager are held by the same person to adopt riskier real earnings management practices.

To test this conjecture, we measure the distribution of decision-making power (*DUAL*) using whether the positions of chairman and general manager are held by the same person. If these positions are combined, *DUAL* is equal to 1, and otherwise 0. The regression results are presented in Table 11. The results show that the coefficient on *Treat*Post* in Column (1) is –0.016 and is significant at the 5 % level. The coefficient on *Treat*Post* in Column (3) is –0.004 but is not significant. The coefficient on *Treat*Post* in Column (2) is –0.006, which is significant at the 5 % level. The coefficient on *Treat*Post* in Column (4) is 0.012, which is significant at the 5 % level. These regression results pass the test of the difference in coefficients between groups, validating our conjecture.

Table 11
Distribution of decision-making power.

Variable	(1) *AbsDA* *DUAL* = 1	(2) *AbsDA* *DUAL* = 0	(3) *AbsREM* *DUAL* = 1	(4) *AbsREM* *DUAL* = 0
*Treat*Post*	–0.016**	–0.006**	–0.004	0.012**
	(–2.331)	(–2.346)	(–0.360)	(2.554)
Size	–0.002	0.002	0.023***	0.018***
	(–0.435)	(1.224)	(3.021)	(4.483)
Maturity	0.006	0.000	–0.001	0.001
	(1.088)	(0.183)	(–0.093)	(0.250)
ATO	0.019***	–0.002	0.139***	0.093***
	(2.708)	(–0.726)	(5.998)	(10.993)
REC	0.087***	0.041***	–0.147**	–0.053*
	(2.960)	(2.708)	(–2.247)	(–1.726)
Board	–0.001	–0.010	–0.029	0.002
	(–0.100)	(–1.563)	(–1.143)	(0.153)
Indep	0.026	–0.030	–0.122*	–0.056
	(0.643)	(–1.628)	(–1.659)	(–1.553)
BM	–0.002	–0.003***	–0.009*	–0.007***
	(–0.581)	(–4.278)	(–1.875)	(–4.654)
Big4	–0.009	–0.003	–0.042	–0.005
	(–0.785)	(–0.556)	(–1.482)	(–0.444)
HHI	–0.016	0.001	–0.032*	–0.008
	(–1.537)	(0.142)	(–1.727)	(–0.803)
Constant	0.117	0.071*	–0.251	–0.250***
	(1.213)	(1.743)	(–1.464)	(–2.783)
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
N	6,972	20,135	6,972	20,135
R-squared	0.099	0.078	0.065	0.045
Chow test p-value		0.0000		0.0000

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

5. Mechanism analysis

AIOI is an important initiative aimed at strengthening SOE supervision, regulating business behavior and preventing the loss of state-owned assets. As previously discussed, we posit that AIOI enhances management supervision and monitoring efficiency by strengthening the external supervision environment and refining the internal organizational environment, thereby influencing SOEs' earnings management strategies. In this section, we examine these two channels: strengthening the external supervision environment through capital market concerns and refining the internal organizational environment through internal control quality. In this way, we can better explore and understand the mechanisms of AIOI and the choice of corporate earnings management strategies, thus revealing the underlying black box.

5.1. External supervision environment

To assess whether AIOI enhances the external supervision environment, we construct the following model for testing:

$$Sup_{i,t} = \beta_0 + \beta_1 Treat * Post + \beta_2 Controls_{i.t} + Year + Firm + \varepsilon_{i,t} \tag{9}$$

In the model, the explanatory variable is the external supervision environment (*Sup*), which is measured using capital market attention (*Anal* and *Resr*). Analyst focus (*Anal*) is equal to the natural logarithm of the number of analysts (teams) following the company in a given year plus 1. Research concern (*Resr*) is equal to the natural logarithm of the number of research reports on the company in a given year plus 1. The key explanatory variable is the interaction term *Treat**Post*. If the value of β_1 is significant and positive, it will confirm that AIOI significantly increases capital market attention, thereby strengthening the external supervision environment. The regression results are presented in Table 12. The coefficients on *Treat**Post* are all positive and significant at least at the 1 % level, indicating that the mechanism of strengthening the external supervision environment is confirmed.

5.2. Internal organizational environment

To investigate whether AIOI enhances the internal organizational environment of SOEs, we construct the following model for testing:

$$IC_{i,t} = \beta_0 + \beta_1 Treat * Post + \beta_2 Controls_{i.t} + Year + Firm + \varepsilon_{i,t} \tag{10}$$

In the model, the explanatory variable is the quality of internal control (*IC*), measured using the internal control index of the Debo Database divided by 100 to ensure the readability of the coefficients. The key explanatory variable is *Treat**Post*. If the value of β_1 is significant and positive, it will confirm that AIOI significantly improves the quality of internal control in SOEs, thereby enhancing the internal organizational environment of SOEs. The regression results are presented in Table 13. Column (1) shows the results of the difference-in-differences regression without control variables, controlling only for firm and year fixed effects. The coefficient on *Treat**Post* is 0.278 and is significant at the 1 % level in Column (1). In Column (2), which includes the control variables and year and individual fixed effects, the coefficient on *Treat**Post* is 0.397 and is significant at the 1 % level. These results indicate that the mechanism of improvement of the internal organizational environment is confirmed.

6. Analysis of economic consequences

SOEs constitute an important material and political foundation of socialism with Chinese characteristics, as well as an important pillar of national economic and social development. They are responsible for preventing the loss of state-owned assets and for ensuring their preservation and appreciation. The previous section proves that AIOI can significantly reduce SOEs' accrual-based earnings management. However, we must also recognize that AIOI leads SOEs to engage in real earnings management. Therefore, the implementation of

Table 12
External supervision environment.

Variable	(1) *Anal*	(2) *Resr*
*Treat*Post*	0.085***	0.110***
	(2.762)	(2.813)
Size	0.550***	0.694***
	(25.489)	(25.226)
Maturity	–0.031	–0.043
	(–1.531)	(–1.614)
ATO	0.348***	0.456***
	(9.207)	(9.502)
REC	0.068	0.077
	(0.384)	(0.352)
Board	0.095	0.170*
	(1.242)	(1.777)
Indep	0.112	0.178
	(0.512)	(0.647)
Dual	0.060***	0.076***
	(2.635)	(2.660)
BM	–0.194***	–0.250***
	(–13.143)	(–13.014)
Big4	0.026	0.017
	(0.443)	(0.230)
HHI	0.155***	0.208***
	(2.800)	(2.923)
Constant	–10.213***	–13.248***
	(–20.306)	(–20.736)
Firm FE	YES	YES
Year FE	YES	YES
N	19,318	19,420
R-squared	0.184	0.165

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

AIOI is a double-edged sword. It therefore remains unclear whether AIOI can ultimately achieve the goal of improving SOEs' economic efficiency, which needs to be further tested.

To examine the impact of AIOI on firm value, we use *TobinQ* to measure firm value. Firm value is used as the explanatory variable, while the interaction term between earnings management and AIOI is used as the explanatory variable. The regression results are presented in Table 14. In Columns (1) and (2), the coefficient of *AbsDA*Treat*Post* on $TobinQ_{t+1}$ is not significant, while the coefficient of *AbsDA*Treat*Post* on $TobinQ_{t+2}$ is 1.519, which is statistically significant at the 5 % level. In Columns (3) and (4), the coefficient of *AbsREM*Treat*Post* on $TobinQ_{t+1}$ is not significant, while the coefficient of *AbsREM*Treat*Post* on $TobinQ_{t+2}$ is –0.675, which is statistically significant at the 10 % level.

The total effect is shown in Columns (5) and (6). Overall, the coefficients of *AbsDA*Treat*Post* and *AbsREM*Treat*Post* on $TobinQ_{t+1}$ are not significant, while the coefficients of *AbsDA*Treat*Post* and *AbsREM*Treat*Post* on $TobinQ_{t+2}$ are significant and positive. These results indicate that AIOI can generally enhance the value of SOEs and promote the preservation and appreciation of state-owned capital. Therefore, the positive effects of AIOI still dominate. However, it is necessary to pay attention to the risky behaviors that may be adopted by SOEs during the implementation of AIOI, to further improve the design of the accountability system.

Table 13
Internal organizational environment.

Variable	*IC*	
	(1)	(2)
*Treat*Post*	0.278***	0.397***
	(6.927)	(9.966)
Size		0.420***
		(14.259)
Maturity		–0.012
		(–0.360)
ATO		0.687***
		(12.520)
REC		–0.134
		(–0.587)
Board		0.066
		(0.561)
Indep		0.523
		(1.632)
Dual		0.067**
		(2.094)
BM		–0.130***
		(–9.441)
Big4		0.004
		(0.044)
HHI		0.293***
		(3.698)
Constant	7.152***	–2.952***
	(274.152)	(–4.269)
Firm FE	YES	YES
Year FE	YES	YES
N	27,107	27,107
R-squared	0.061	0.096

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

7. Conclusion

The Opinions issued by the State Council in 2016 introduced a groundbreaking concept: "the implementation of a system of lifelong accountability for major decisions," marking a significant shift in the governance of SOEs. This system imposes strict constraints on the operating and investment decisions of SOEs, generating both *ex-post* punishment and *ex-ante* deterrent effects on management. Leveraging the quasi-natural experiment of the introduction of the Opinions, we use a difference-in-differences approach to investigate the impact on SOEs' earnings management. Our analysis reveals that compared with non-SOEs, AIOI effectively reduces accrual-based earnings management in SOEs, while encouraging them to engage in real earnings management practices. Our robustness tests, including placebo tests and PSM-DID, strengthen the causality of our findings. Further cross-sectional analyses demonstrate that the influence of AIOI on earnings management varies, with greater effects observed in SOEs with weaker digital transformation and decentralized decision-making power. The mechanism analysis reveals that AIOI reshapes SOEs' earnings management behavior by improving the external supervision environment and the internal organizational environment, thus prompting strategic changes in their approach to earnings management. Finally, the analysis of the economic consequences shows that AIOI can ultimately enhance the overall value of SOEs.

This study contributes to the literature on the microeconomic effects of AIOI and the determinants of earnings management strategies. Unlike prior research, which often overlooks the role of accountability systems in

Table 14
Analysis of economic consequences.

Variable	(1) $TobinQ_{t+1}$	(2) $TobinQ_{t+2}$	(3) $TobinQ_{t+1}$	(4) $TobinQ_{t+2}$	(5) $TobinQ_{t+1}$	(6) $TobinQ_{t+2}$
*AbsDA*Treat*Post*	0.500	1.519**			0.797	1.850**
	(0.428)	(2.121)			(0.663)	(2.549)
AbsDA	1.964***	0.365			1.516***	0.194
	(4.222)	(1.069)			(3.126)	(0.564)
*AbsREM*Treat*Post*			–0.473	–0.675*	–0.603	–0.929**
			(–1.003)	(–1.811)	(–1.227)	(–2.454)
AbsREM			1.371***	0.525**	1.193***	0.503**
			(5.880)	(2.222)	(4.894)	(2.111)
*Treat*Post*	–0.611***	–0.500***	–0.543***	–0.319***	–0.564***	–0.405***
	(–5.659)	(–4.757)	(–4.961)	(–2.982)	(–4.887)	(–3.626)
Size	–0.169*	0.087	–0.190*	0.087	–0.187*	0.084
	(–1.722)	(0.870)	(–1.917)	(0.856)	(–1.895)	(0.824)
Maturity	0.055	0.028	0.057	0.028	0.056	0.028
	(0.714)	(0.297)	(0.736)	(0.295)	(0.732)	(0.295)
ATO	0.814***	0.449***	0.685***	0.396**	0.706***	0.415**
	(5.524)	(2.804)	(4.743)	(2.471)	(4.828)	(2.573)
REC	–1.141	–0.591	–0.961	–0.522	–1.033	–0.553
	(–1.544)	(–0.682)	(–1.312)	(–0.604)	(–1.402)	(–0.639)
Board	0.031	0.061	0.031	0.060	0.040	0.063
	(0.121)	(0.223)	(0.118)	(0.219)	(0.156)	(0.229)
Indep	–0.517	0.627	–0.456	0.630	–0.435	0.667
	(–0.676)	(0.850)	(–0.593)	(0.852)	(–0.569)	(0.904)
Dual	–0.073	–0.043	–0.071	–0.042	–0.075	–0.043
	(–0.922)	(–0.659)	(–0.907)	(–0.655)	(–0.962)	(–0.662)
BM	0.036	–0.015	0.039	–0.014	0.042	–0.013
	(1.052)	(–0.438)	(1.136)	(–0.417)	(1.224)	(–0.387)
Big4	0.299	0.264	0.322	0.267	0.321	0.276
	(1.221)	(1.012)	(1.312)	(1.015)	(1.304)	(1.053)
HHI	0.034	0.062	0.061	0.072	0.053	0.066
	(0.172)	(0.299)	(0.309)	(0.348)	(0.271)	(0.319)
Constant	8.113***	2.424	8.502***	2.392	8.301***	2.431
	(3.604)	(1.030)	(3.750)	(1.011)	(3.689)	(1.026)
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
N	21,360	17,841	21,360	17,841	21,360	17,841
R-squared	0.134	0.136	0.135	0.136	0.137	0.137

Notes: The t-values are reported in parentheses; ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

shaping corporate disclosure decisions, we specifically examine the impact of AIOI on SOEs' earnings management, simultaneously considering accrual-based and real earnings management. Our findings offer valuable insights with multiple policy implications, providing empirical support for governments to evaluate and enhance accountability systems at the micro level. Moreover, our study offers policy guidance for effectively curbing management irregularities, protecting state-owned assets and fostering the development of the capital market. However, our results highlight a concerning transfer effect between accrual-based earnings management and real earnings management in SOEs under pressure from the accountability system. The question of how to better regulate conflicts of interest and short-sighted behavior arising from principal–agent problems is a common concern in academic and transactional circles.

CRediT authorship contribution statement

Xuena Liu: Conceptualization, Methodology, Writing – review & editing. **Jiemei Liu:** Conceptualization, Data curation, Software. **Lu Pan:** Conceptualization, Data curation, Writing – original draft.

Funding

This work was supported by "the Fundamental Research Funds for the Central Universities", Zhongnan University of Economics and Law (Grant Nos. 202411117).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmed, A.S., Duellman, S., Grady, M., 2022. Political connections and the trade-off between real and accrual-based earnings management. Contemp. Account. Res. 39 (4), 2730–2757. https://doi.org/10.1111/1911-3846.12811.

Alhadab, M., Clacher, I., Keasey, K., 2014. Real and accrual earnings management and IPO failure risk. Account. Bus. Res. 45 (1), 55–92. https://doi.org/10.1080/00014788.2014.969187.

Baker, M., Wurgler, J., 2004. A catering theory of dividends. J. Financ. 59 (3), 1125–1165. https://doi.org/10.1111/j.1540-6261.2004.00658.x.

Becker, G.S., 1968. Crime and punishment: an economic approach. J. Polit. Econ. 76 (2), 169–217. https://doi.org/10.1086/259394.

Bereskin, F.L., Hsu, P., Rotenberg, W., 2018. The real effects of real earnings management: evidence from innovation. Contemp. Account. Res. 35 (1), 525–557. https://doi.org/10.1111/1911-3846.12376.

Beuselinck, C., Cascino, S., Deloof, M., Vanstraelen, A., 2018. Earnings management within multinational corporations. Account. Rev. 94 (4), 45–76. https://doi.org/10.2308/accr-52274.

Cai, W., Cai, X., Wang, Z., Yang, G., 2023. The spillover effect of penalty against peer firm leaders——evidence from earnings management. Financ. Res. Lett. 54, 103701. https://doi.org/10.1016/j.frl.2023.103701.

Chang, K., Kim, Y.S., Li, Y., Park, J.C., 2022. Labor unions and real earnings management. J. Corp. Financ. 75, 102242. https://doi.org/10.1016/j.jcorpfin.2022.102242.

Chen, H., Tian, Z., 2022. Environmental uncertainty, resource orchestration and digital transformation: a fuzzy-set QCA approach. J. Bus. Res. 139, 184–193. https://doi.org/10.1016/j.jbusres.2021.09.048.

Chen, C.C.S., Chou, Y.Y., Wei, P., 2020. Country factors in earnings management of ADR firms. Financ. Res. Lett. 32, 101146. https://doi.org/10.1016/j.frl.2019.04.003.

Chen, G., Wang, M., 2023. Stock market liberalization and earnings management: evidence from the china–Hong Kong stock connects. Financ. Res. Lett. 58, 104417. https://doi.org/10.1016/j.frl.2023.104417.

Chen, X., Cheng, Q., Wang, X., 2015. Does increased board independence reduce earnings management? Evidence from recent regulatory reforms. Rev. Account. Stud. 20 (2), 899–933. https://doi.org/10.1007/s11142-015-9316-0.

Chen, Y.S., Jiang, Y., He, Y.R., 2022. Accountability for illegal operation and investment and state-owned enterprises' risk-taking. Account. Res. 4, 53–70 (in Chinese).

Cheng, Q., Lee, J., Shevlin, T., 2015. Internal governance and real earnings management. Account. Rev. 91 (4), 1051–1085. https://doi.org/10.2308/accr-51275.

Choi, A., Choi, J., Sohn, B.C., 2018. The joint effect of audit quality and legal regimes on the use of real earnings management: international evidence. Contemp. Account. Res. 35 (4), 2225–2257. https://doi.org/10.1111/1911-3846.12370.

Chung, C.Y., Hwang, J.H., Kim, D., Liu, C., 2018. When institutions passively curb earnings management: evidence from the Korean market. Financ. Res. Lett. 25, 274–279. https://doi.org/10.1016/j.frl.2018.02.033.

Chung, H., Sheu, H.J., Wang, J.L., 2009. Do firms' earnings management practices affect their equity liquidity? Financ. Res. Lett. 6 (3), 152–158. https://doi.org/10.1016/j.frl.2009.03.003.

Cohen, D.A., Dey, A., Lys, T.Z., 2008. Real and accrual-based earnings management in the pre- and post-Sarbanes-oxley periods. Account. Rev. 83 (3), 757–787. https://doi.org/10.2308/accr.2008.83.3.757.

Commerford, B.P., Hatfield, R.C., Houston, R.W., 2018. The effect of real earnings management on auditor scrutiny of management's other financial reporting decisions. Account. Rev. 93 (5), 145–163. https://doi.org/10.2308/accr-52032.

Dechow, P.M., Sloan, R.G., Sweeney, A.P., 1995. Detecting earnings management. Account. Rev. 70 (2), 193–225 http://www.jstor.org/stable/248303.

Dimmock, S.G., Feng, F., Zhang, H., 2023. Mutual funds' capital gains lock-in and earnings management. J. Corp. Financ. 80, 102422. https://doi.org/10.1016/j.jcorpfin.2023.102422.

Ege, M., Seidel, T.A., Sterin, M., Wood, D.A., 2022. The influence of management's internal audit experience on earnings management. Contemp. Account. Res. 39 (3), 1834–1870. https://doi.org/10.1111/1911-3846.12770.

Evans, M.E., Houston, R.W., Peters, M.F., Pratt, J.H., 2015. Reporting regulatory environments and earnings management: U.S. and non-U.S. firms using U.S. GAAP or IFRS. Account. Rev. 90 (5), 1969–1994. https://doi.org/10.2308/accr-51008.

Ewert, R., Wagenhofer, A., 2005. Economic effects of tightening accounting standards to restrict earnings management. Account. Rev. 80 (4), 1101–1124. https://doi.org/10.2308/accr.2005.80.4.1101.

Fan, Z., Radhakrishnan, S., Zhang, Y., 2020. Corporate governance and earnings management: evidence from shareholder proposals. Contemp. Account. Res. 38 (2), 1434–1464. https://doi.org/10.1111/1911-3846.12640.

Gao, L., Han, J., Kim, J.B., Pan, Z., 2024. Overlapping institutional ownership along the supply chain and earnings management of supplier firms. J. Corp. Financ. 84, 102520. https://doi.org/10.1016/j.jcorpfin.2023.102520.

Griffin, P.A., Hong, H.A., Liu, Y., Ryou, J.W., 2021. The dark side of CEO social capital: evidence from real earnings management and future operating performance. J. Corp. Financ. 68, 101920. https://doi.org/10.1016/j.jcorpfin.2021.101920.

Guggenmos, R.D., Van der Stede, W.A., 2020. The effects of creative culture on real earnings management. Contemp. Account. Res. 37 (4), 2319–2356. https://doi.org/10.1111/1911-3846.12586.

He, H., Fang, J., 2024. Does the integration between litigation and supervision discipline financial misstatement? China J. Account. Res. 17 (2), 100357. https://doi.org/10.1016/j.cjar.2024.100357.

Huang, S., Roychowdhury, S., Sletten, E., 2019. Does litigation deter or encourage real earnings management? Account. Rev. 95 (3), 251–278. https://doi.org/10.2308/accr-52589.

Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: managerial behavior, agency costs and ownership structure. J. Financ. Econ. 3 (4), 305–360. https://doi.org/10.1016/0304-405x(76)90026-x.

Kim, M.P., Pierce, S.R., Yeung, I., 2021. Why firms announce good news late: earnings management and financial reporting timeliness. Contemp. Account. Res. 38 (4), 2691–2722. https://doi.org/10.1111/1911-3846.12695.

Le, T.D., Trinh, T., 2022. Distracted analysts and earnings management. Financ. Res. Lett. 49, 103038. https://doi.org/10.1016/j.frl.2022.103038.

Li, D., Huang, H., Wang, K., 2022. Information technology infrastructure and earnings management strategy: evidence from a quasi-natural experiment. China J. Account. Stud. 11 (1), 108–133. https://doi.org/10.1080/21697213.2023.2148949.

Licht, A.N., 2002. Accountability and corporate governance. SSRN Electron. J. https://doi.org/10.2139/ssrn.328401.

Lin, K.J., Lu, X., Zhang, J., Zheng, Y., 2020. State-owned enterprises in China: a review of 40 years of research and practice. China J. Account. Res. 13 (1), 31–55. https://doi.org/10.1016/j.cjar.2019.12.001.

Liu, M., Li, C., Wang, S., Li, Q., 2023. Digital transformation, risk-taking, and innovation: evidence from data on listed enterprises in China. J. Innov. Knowl. 8 (1), 100332. https://doi.org/10.1016/j.jik.2023.100332.

Matsumoto, D.A., 2002. Management's incentives to avoid negative earnings surprises. Account. Rev. 77 (3), 483–514. https://doi.org/10.2308/accr.2002.77.3.483.

Ng, J., Wu, H., Zhai, W., Zhao, J., 2021. The effect of shareholder activism on earnings management: evidence from shareholder proposals. J. Corp. Financ. 69, 102014. https://doi.org/10.1016/j.jcorpfin.2021.102014.

Oz, I.O., Yelkenci, T., 2018. Examination of real and accrual earnings management: a cross-country analysis of legal origin under IFRS. Int. Rev. Financ. Anal. 58, 24–37. https://doi.org/10.1016/j.irfa.2018.04.003.

Panicker, V.S., Upadhyayula, R.S., Sivakumar, S., 2022. Internationalization of hybrid state-owned enterprises from emerging markets: institutional investors as enablers. J. Bus. Res. 151, 409–422. https://doi.org/10.1016/j.jbusres.2022.07.018.

Raman, K., Shahrur, H., 2008. Relationship-specific investments and earnings management: evidence on corporate suppliers and customers. Account. Rev. 83 (4), 1041–1081. https://doi.org/10.2308/accr.2008.83.4.1041.

Roychowdhury, S., 2006. Earnings management through real activities manipulation. J. Account. Econ. 42 (3), 335–370. https://doi.org/10.1016/j.jacceco.2006.01.002.

Schillemans, T., Busuioc, M., 2015. Predicting public sector accountability: from agency drift to forum drift. J. Publ. Adm. Res. Theor. 25 (1), 191–215. https://doi.org/10.1093/jopart/muu024.

Shen, J., Deng, X., Hou, Q., 2024. The monkey king wearing a tight band: executive accountability pressure and corporate investment. Financ. Res. Lett. 68, 105979. https://doi.org/10.1016/j.frl.2024.105979.

Song, B., Chung, H., Kim, B.J., Sonu, C.H., 2023. Do business trainings for audit committees matter in organizations? focusing on earnings management. Financ. Res. Lett. 51, 103423. https://doi.org/10.1016/j.frl.2022.103423.

Sun, Q., Zhang, A., Li, J., 2005. A study of optimal state shares in mixed oligopoly: implications for SOE reform and foreign competition. China Econ. Rev. 16 (1), 1–27. https://doi.org/10.1016/j.chieco.2004.06.009.

Wang, S., Chen, X., Wang, Q., 2025. Is corporate digital transformation counter-cyclical? China J. Account. Res. 18 (1), 100401. https://doi.org/10.1016/j.cjar.2024.100401.

Wu, C., Xu, L., Xin, Y., 2023. The state capital investing and operating company pilot reform and SOE bailouts. China J. Account. Res. 16 (2), 100302. https://doi.org/10.1016/j.cjar.2023.100302.

Wu, Y.J., Chen, W., Yan, H., 2024. The accountability system for operation and investment and audit quality of state-owned enterprises. Int. Rev. Financ. Anal., 103680 https://doi.org/10.1016/j.irfa.2024.103680.

Xin, Y., Song, P.X., Xu, L.P., Teng, F., 2022. The accountability system for operation and investment and the standardized operation of state-owned enterprises: an empirical study based on managers' violation behaviors. J. Manage. World 38 (12), 199–221. https://doi.org/10.19744/j.cnki.11-1235/f.2022.0180 (in Chinese).

Zang, A.Y., 2012. Evidence on the trade-off between real activities manipulation and accrual-based earnings management. Account. Rev. 87 (2), 675–703. https://doi.org/10.2308/accr-10196.

Zheng, D., Xu, Y., Wenren, Y., 2024. Compliance management and investment efficiency in state-owned enterprises: evidence from China. China J. Account. Res. 17 (2), 100358. https://doi.org/10.1016/j.cjar.2024.100358.

China Journal of Accounting Research 18 (2025) 100422

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Communication media and linguistic concreteness: evidence from firm site visits in China

Songlian Tang, Jiangnan Geng *

Glorious Sun School of Business and Management, Donghua University, 1882 West Yan'an Road, Shanghai 200051, China

ARTICLE INFO

Article history:
Received 29 May 2024
Accepted 12 April 2025
Available online 22 May 2025

Keywords:
Institutional investor visits
Online communication
Linguistic concreteness
Analyst forecasts

ABSTRACT

The digital transformation of communication media has been widely adopted and has reshaped institutional investor interactions. We examine how the linguistic concreteness of investors' questions adapts to online media and face-to-face communication, using textual analysis of investor–company Q&A transcripts from Chinese listed firms. We find that online media environments induce investors to enhance linguistic precision, to compensate for reduced nonverbal cues and social presence. This mechanism is moderated by sender characteristics, message attributes and response concreteness patterns. Such adjustments counteract media-induced information loss, evidenced by improved analyst forecast accuracy. Our study represents a practical application of Lasswell's "5W" communication model in the context of corporate information disclosure and expands research on the impact of digital technology on communication media, offering insights into the balance between sender intent and media characteristics.

1. Introduction

Digital technology is driving the evolution of communication media, freeing information from its reliance on physical forms and extending interactions from the field to the online space. Institutional investor visits, serving as the primary channel of communication between listed firm management and investors, has become the cornerstone of investor relations activities. In 2022, the China Securities Regulatory Commission revised its *Guidelines for Investor Relations Management*, encouraging listed firms to engage with investors through diverse channels and platforms. Consequently, online communication in investor relations increased from 4.62 % in 2019 to 32.5 % in 2021, creating a situation where online and site visits coexist.

* Corresponding author.
E-mail addresses: tangsonglian@dhu.edu.cn (S. Tang), jenna_k@163.com (J. Geng).

https://doi.org/10.1016/j.cjar.2025.100422
1755-3091/

Lasswell's "5W" model divides communication into five components—sender, message, medium, receiver and effect—with the medium linking the other elements (Zhang, 2004). Media richness theory further asserts that different media vary in their ability to convey information (Daft and Lengel, 1986; Carlson and Robert, 1999). For example, face-to-face communication can convey nonverbal cues such as tone, facial expressions and gestures, which help listeners decode the sender's thoughts and emotions, facilitate sustained mutual attention and regulate turn-taking in conversation (Tam and Tian, 2023). However, this raises a key question: How can senders balance their desire to share information with the medium's limitations to ensure effective communication?

During institutional site visits, investors gain first-hand insights by visiting production facilities and observing operational details (Cheng et al., 2016). They also capture additional hidden information through executives' nonverbal cues such as facial expressions, gestures and gazes. In contrast, online institutional visits rely on video conferencing or teleconference Q&A sessions, where investors obtain information remotely. Compared with face-to-face communication, the absence of nonverbal cues in online interactions (Hiltz et al., 1986; Kimura et al., 2021) reduces the precision of analysts' earnings forecasts (Kang and Li, 2024).

Language, as a carrier of information, focuses more on content, paying less attention to the mode of expression (Bozeman and Kacmar, 1997). Linguistic concreteness is an attribute of language characterized by detailed and descriptive expressions that include specific information and situational context (Langacker, 1987; Hansen and Wänke, 2010). For receivers, such concreteness may activate dual coding systems (Paivio, 1991), prompting investors to react more proactively in the face of information asymmetry and uncertainty (Larrimore et al., 2011; Toma and D'Angelo, 2015). This raises the following question: In the context of institutional visits, do investors adjust their language concreteness to mitigate the adverse effects of online media?

Leveraging the coexistence of online and site visits, we develop an index of investor language concreteness through textual analysis of investor relations event disclosure records of Shenzhen-listed companies (2012–2021). We investigate whether enhancing language concreteness can mitigate the disadvantages of online communication. Our main findings are as follows. First, investors use more concrete language in online communications than during in-person site visits, a pattern consistently confirmed by a series of robustness checks. Second, our heterogeneity analysis reveals that language concreteness is influenced by sender, message and receiver characteristics, varying across factors such as the privacy of institutional visits, information uniqueness, the characteristics of board secretaries, management position and managerial language concreteness. Third, using specific language in online visits reduces bias and optimism in analysts' forecasts, thus mitigating the adverse impact of online media on communication effectiveness.

Our study contributes to the literature in the following ways. First, we advance research on the evolution of digital communication by using linguistic analysis to differentiate online and offline interactions and examine the interplay between communicative intent and medium constraints. Second, we link communication theory to corporate disclosure practices by applying Lasswell's "5W" model, extending the literature from the effects of media on governance structures to their impacts on communicators' willingness to engage. Finally, we enrich research on language attributes by developing a text-based concreteness index from transcripts of investor–management Q&As, thereby expanding the traditional focus on content/emotional attributes (Hansen and Wänke, 2010; Pan et al., 2018) to include dimensions of linguistic expression.

The remainder of this paper is organized as follows. Section 2 reviews the literature and develops the hypotheses. Section 3 outlines the research design. Section 4 presents our empirical analysis. Section 5 discusses our heterogeneity analysis. Section 6 analyzes the economic effects. Finally, Section 7 concludes the paper with a summary of our research findings and policy recommendations.

2. Literature review and hypothesis development

2.1. Literature review

2.1.1. Communication media

Harold Lasswell's "5W" model breaks down the communication process into five components, namely sender, message, medium, receiver and effect, with the medium serving as the critical link between them and influencing the effect of communication. In 1986, Daft and Lengel proposed media richness theory to examine how different communication media affect outcomes. In addition, research shows that face-to-face communication generates greater trust (Nilsson and Mattes, 2015), stronger motivation (Storper and Venables, 2004), greater satisfaction (Purvanova, 2014) and more positive perceptions of others.

Previous studies generally compare communication media by focusing on nonverbal cues and social presence. First, face-to-face communication can transmit nonverbal signals such as tone, facial expressions and gestures, which help listeners interpret the content and provide the greatest amount of information (Williams, 1977). Second, face-to-face interactions generate the highest level of social presence, enabling participants to perceive each other's presence, which effectively alleviates tension and stress (Parker et al., 1976). These conclusions apply to both online education and virtual meetings. Social interaction patterns between teachers and students differ between face-to-face and online classes (Szeto and Cheng, 2016), and participants in online meetings often struggle to fully convey verbal and nonverbal signals (Kimura et al., 2021; Tam and Tian, 2023).

In recent years, online communication has been widely adopted in corporate governance and information disclosure. Online communication not only saves time and reduces travel costs but has also become an effective corporate governance tool by enhancing the reach of meetings and reducing hierarchical pressures in communication. First, in terms of environmental sustainability, online board meetings reduce environmental pollution associated with air travel required for face-to-face meetings (Kruger and Chowers, 2020). Second, in terms of increasing meeting reach, Gao et al. (2020) find that online shareholder meetings can significantly boost shareholder participation. Third, in terms of reducing social hierarchy pressure, Cai et al. (2023) observe that during online board meetings, directors are more likely to raise objections regarding agency conflicts.

Despite these advantages, prior studies comparing face-to-face and online communication have not reached a clear consensus, largely because previous research typically overlooks the transmitter's willingness to communicate. Moreover, many studies focus on scenarios in which the communication process is not fully disclosed, thus failing to capture the sender's willingness to engage in either mode.

2.1.2. *Institutional investor visits*

Institutional investor visits serve as an interactive communication channel between outside investors and corporate insiders and can be classified into site visits and online visits. Extensive research on investors' site visits focuses on several areas, including their impact on stock market pricing efficiency and market information content (Cheng et al., 2016; Li et al., 2022; Bowen et al., 2023; Cheng et al., 2019; Xiang and Yang, 2022); their influence on institutional investors' behavior in the capital market (Tan and Cui, 2015; Xiao et al., 2017; Yan et al., 2024) and funds' excess returns (Liu et al., 2017; Xu et al., 2021; Yue et al., 2023); and their governance effects on listed companies (Tan and Lin, 2016; Cheng et al., 2017; Lu et al., 2019; Bu and Sun, 2020; Peng et al., 2022; Cao et al., 2022; Song and Xian, 2024; Wang et al., 2024; Zhang et al., 2024).

With the advent of digital technology, investors' online visits have become increasingly common. Li and Zhao (2021) demonstrate that such visits can enhance firms' total factor productivity. Moreover, Chen et al. (2021) find that compared with analysts who do not participate, those who engage in online visits make more accurate earnings forecasts and trigger a stronger market reaction. However, some scholars argue that the lack of on-site observation during investors' online visits may reduce the accuracy of analysts' earnings per share (EPS) forecasts relative to those based on investors' site visits (Kang and Li, 2024).

2.1.3. *Language concreteness*

Linguistic concreteness refers to detailed descriptive language that conveys precise information and context (Langacker, 1987; Hansen and Wänke, 2010), indicating the amount of informational detail and context that a language expression provides. Dual coding theory holds that the human cognitive system comprises both verbal and nonverbal coding systems and that concrete language can activate both systems simultaneously, resulting in a doubled short-term memory effect compared with a single coding system (Paivio, 1991; Huang, 2018). Consequently, audiences can rapidly capture the information contained in a specific language (Estes, 1982; Sadoski, 2001), thus enhancing the efficiency of information processing (McClelland and Rumelhart, 2020), which in turn improves the sender's credibility and trustworthiness and ultimately bolsters the persuasiveness of the communication.

Pan et al. (2018) conceptualize language concreteness from a part-of-speech perspective, arguing that concrete language is characterized by the use of verbs, numbers and past-tense words. Elliott et al. (2015) find that when investors perceive higher risk in a company, they respond more proactively to management's concrete language. Specifically, management's use of precise language in prospectuses boosts investors' acceptance of the company's valuation and increases their willingness to invest, while its use in the management discussion and analysis sections of annual reports reduces stock price synchronization by lowering information processing costs and enhancing credibility (Zhao et al., 2022).

2.2. Hypothesis development

The process of information dissemination involves five components, namely sender, message, medium, receiver and effect, with the medium connecting the sender, message and receiver. Comparative studies reveal that face-to-face communication is more effective in eliciting behavioral changes than video conferencing (Nilsson and Mattes, 2015). Specifically, face-to-face communication can convey nonverbal cues such as tone, facial expressions and gestures, which provide the richest information (Williams, 1977), and offer the highest level of social presence, effectively alleviating tension and stress (Parker et al., 1976).

Institutional visits provide investors and company insiders with a channel for interaction. With the rise of the digital economy, online visits have become increasingly common, alongside traditional investor site visits (Li and Zhao, 2021). This evolution is linked to the exchange of information through investor–executive Q&A dialogues. Regardless of the medium of communication, investors systematically demonstrate multilevel information needs, ranging from operational specificities to strategic positioning. Investor questions cover a wide range of topics (Han et al., 2018), and company insiders not only answer these questions (Solomon and Soltes, 2015) but also discuss the company in depth within the bounds of regulatory compliance.

During site visits, investors can extract multifaceted information. They have the opportunity to tour factories and production facilities to learn about the latest operational conditions and can capture real-time cues such as management's voice, expressions and gestures, which together create a strong sense of presence in the interaction. Although digital platforms enhance connectivity through efficient and cost-effective communication, they fundamentally lack immersive situational awareness. This technological gap creates cognitive barriers in virtual interactions, where participants must compensate for decreased social presence and lack of nonverbal cues (Kimura et al., 2021). Fundamentally, as evidenced by Kang and Li (2024), online visits demonstrably impair financial analysts' forecast accuracy compared with physical site evaluations.

Language is an essential vehicle for the transfer of information and its concreteness is crucial for effective communication. Linguistic concreteness enhances both the efficiency of information processing and its persuasiveness. Dual coding theory holds that concrete language simultaneously activates both verbal and nonverbal systems (Paivio, 1991), enabling receivers to process information more efficiently and accurately. Additionally, concrete language increases communicative persuasiveness (Toma and D'Angelo, 2015), yielding more favorable responses under conditions of high information asymmetry and uncertainty (Larrimore et al., 2011; Pan et al., 2018; Zhao et al., 2022). Online visits require investors to strategically calibrate linguistic precision, compensating for the structural constraints of digital communication channels while optimizing the fidelity of information transmission.

In summary, investors leverage concrete linguistic expressions during institutional visits to compensate for the lack of nonverbal cues and reduced social presence inherent in online communication. Therefore, we propose our main hypothesis:

Hypothesis 1 (H1): Compared with site visits, investors' language during online visits is more concrete.

3. Data and methodology

3.1. Sample and data

In 2012, the Shenzhen Stock Exchange (SZSE) required all listed firms to disclose detailed information regarding site visits within two trading days of each visit (Bowen et al., 2018). This disclosure includes all pertinent details, such as the date of the visit, the venue, the list of visiting institutions and the questions and responses exchanged during the Q&A sessions. We select all SZSE-listed companies from 2012 to 2021 as our initial sample[1] and apply the following processing steps: (1) exclusion of financial companies; (2) exclusion of ST companies; (3) elimination of observations with missing data; and (4) aggregation of investor–company Q&A session data to the research level. After these steps, there remain 50,934 research-level observations in the sample. Table 1 presents the sample selection process. The sample data come from the China Stock Market and Accounting Research database and all continuous variables are winsorized at the 1 % and 99 % quantiles to mitigate the effect of extreme values.

3.2. Variable definitions

Investor language concreteness (*Q_CON*): This variable is constructed following Pan et al. (2018) through a three-step process: thesaurus construction, Q&A-level indicator calculation and research-level aggregation.

Thesaurus construction: Concrete language is defined as including verbs, numbers and past-tense words, while non-concrete language includes adjectives, non-specific quantifiers and future-oriented words (Semin and Fiedler, 1988; Elliott et al., 2015; Snefjella and Kuperman, 2015). First, we collect texts by downloading investor relations event disclosure records of SZSE-listed companies (2012–2021) and randomly selecting 100 Q&A minutes per year (for a total of 1,000 min). Next, nine accounting graduate students are divided into three groups (A, B and C), with groups A and B independently extracting vocabulary meeting the concreteness criteria, and then performing cross-validation by exchanging samples and re-extracting the vocabulary. Finally, group C independently reviews the results, retaining only the vocabulary unanimously approved by all groups. The final thesaurus comprises six categories and a total of 60,826 words, including 47,396 verbs, 21 quantifiers, 263 past-tense words, 12,633 adjectives, 202 non-specific quantifiers and 311 future-oriented words.

Q&A-level calculation: Using the constructed thesaurus and a lexicon-based method in Python, we count the frequency of the six word types in each Q&A. The language concreteness of each Q&A is calculated as follows: Language Concreteness = (Frequency of concrete words) – (Frequency of non-concrete words).

Investor visit-level aggregation: We compute the average language concreteness of the Q&A sessions for each visit and apply a logarithmic transformation to generate the final investor language concreteness index.

Medium of communication (*MediaMode*): The dummy variable for communication medium is coded 1 if the location in the investor relations interaction records contains keywords indicating digital channels (e.g., phone, online, video conference, Tencent, livestream, web-based, scheduled virtual meeting, remote or other equivalent terms), indicating investors' online visits. All other cases are coded 0 to represent investors' site visits.

Control variables: Following Pan et al. (2018), Li and Zhao (2021) and Yao et al. (2021), we control for firm-level financial variables, namely firm size (*Size*), gearing ratio (*Lev*), return on equity (*ROE*), sales growth (*Growth*) and market capitalization-to-book value ratio (*MB*), and corporate governance variables, namely firm age (*Age*), proportion of independent directors (*IndDire*), board size (*Board*), proportion of institutional investors (*InsInve*), ownership type (*SOE*), dual positions (*Dual*), foreign background (*Foreign*), Big 4 audit (*Big4*) and cross-listing (*Cross*). Additionally, at the visit level, we include controls for a Q&A's emotional tone (*Tone*), total number of characters (*Len*) and number of Q&A sessions (*Vn*), along with year and industry fixed effects. All variables are defined in Table 2.

3.3. Research design

Research model (1) is constructed to explore the relationship between the medium of communication and the linguistic concreteness of investors:

$$Q\_CON_{i,t} = \beta_0 + \beta_1 MediaMode_{i,t} + \beta_2 Controls_{i,t} + YearFE + FirmFE + \varepsilon_{i,t} \quad (1)$$

[1] In 2012, the SZSE began requiring listed companies to disclose information related to investor visits, while the Shanghai Stock Exchange only requires listed companies to submit a summary report after such visits without public disclosure.

Table 1
Sample selection process.

Year	Initial sample (data at the Q&A level)	Deletion of financial companies	Deletion of ST companies	Deletion of missing data	Effective sample (data at the Q&A level)	Sample (data for this research)
2012	18,450	(1,454)	0	0	16,996	1,981
2013	54,095	(2,691)	0	(153)	51,251	6,202
2014	58,350	(4,191)	0	(86)	54,073	6,705
2015	45,463	(3,656)	0	(343)	41,464	5,510
2016	48,545	(3,261)	0	(271)	45,013	5,978
2017	42,827	(3,004)	(6)	(199)	39,618	5,230
2018	36,747	(1,664)	(17)	(103)	34,963	4,614
2019	34,758	(1,503)	(36)	(91)	33,128	4,326
2020	36,245	(886)	(13)	(94)	35,252	4,609
2021	47,969	(2,019)	0	(10)	45,940	5,779
Total	423,449	(24,329)	(72)	(1,350)	397,698	50,934

In model (1), $Q\_CON_{i,t}$ is the investor language concreteness of firm *i*. $MediaMode_{i,t}$ is the medium of communication, indicating whether a visit is carried out online or not, and $\varepsilon_{i,t}$ is the random perturbation term. If β_1 is significant and positive, it indicates that investors' questions are more concrete during online visits, i.e., investors will optimize the information dissemination effect by improving the concreteness of language in online communication. Meanwhile, to enhance the reliability of our empirical results, we control for year and firm fixed effects and cluster robust standard errors at the investor visit level.

4. Empirical results

4.1. Descriptive statistics

Table 3 presents the descriptive statistics. The mean value of *Q_CON* in Panel A is 0.889, which implies that on average, each investor's question contains 1.433 ($= e^{0.889} - 1$) words, reflecting their level of linguistic concreteness. The mean value of *MediaMode* is 0.077, indicating that approximately 8 % of the investor visits in our sample are conducted online.

The control variables are consistent with those reported in prior studies. For example, the mean of *IndDire* is 0.376, suggesting that independent directors account for approximately one third of the board of Chinese listed companies (Bowen et al., 2018). The mean of *Tone* is 0.699, reflecting an overall positive tone in investor–management communications. Moreover, Panel B shows that since 2019, the mean of *MediaMode* has increased over the years, highlighting the new coexistence of online and site visits in the communication of investor and listed firms.

4.2. Univariate test

Table 4 presents the univariate test results comparing online visits and site visits. The sample of online visits has a mean *Q_CON* of 0.905, which is significantly higher (at the 10 % level) than the 0.888 observed in the sample of site visits. These findings indicate that investors use more concrete language during online visits, providing initial support for H1.

4.3. Correlation analysis

The results of the correlation analysis for the main variables are shown in Table 5. The Pearson correlation coefficient between *MediaMode* and *Q_CON* is 0.008 and is significant at the 10 % level, indicating that investors' linguistic concreteness is significantly higher during online visits, providing preliminary support for H1. Regarding the results for the control variables, *MediaMode* shows a positive correlation with *Size*, *ROE*, *Age*,

Table 2
Variable definitions.

Variable	Meaning	Definition
Q_CON	Investor language concreteness	Measure of the concreteness of investors' questions, referring to Pan et al. (2018)
MediaMode	Medium of communication	Dummy variable, online visits are assigned a value of 1, and otherwise 0
Size	Size	Natural logarithm of a firm's total assets at the beginning of the year
Lev	Leverage ratio	Ratio of total liabilities to total assets of a company
ROE	Return on net assets	Ratio of net profit to net assets of a company
Growth	Sales revenue growth rate	(Current year operating income – Prior year operating income) / Prior year operating income
MB	Ratio of market value to book value	Ratio of the market value to the book value of a company's equity capital
Age	Firm age	Number of years a firm has been listed
IndDire	Percentage of independent directors	Ratio of the number of independent directors to the size of the board
Board	Board size	Total number of board members of a company
InsInve	Institutional investor holdings	Ratio of shares held by institutional investors to a company's total shares outstanding
SOE	State ownership	Dummy variable, assigned a value of l if the controlling shareholder of a company is the state, and 0 otherwise
Dual	Dual positions	Dummy variable, assigned a value of 1 when the CEO and chair of the board are the same person, and otherwise 0
Foreign	Overseas background of executives	Dummy variable, assigned a value of l if an executive has studied or worked abroad, and 0 otherwise
Big4	Whether a firm is audited by the "Big 4"	Dummy variable, assigned a value of 1 if a company is audited by a Big 4 firm, and otherwise 0
Cross	Indicator of cross-listed firms	Dummy variable, assigned a value of l if a firm is cross-listed, and 0 otherwise
Tone	Emotional tone of Q&As	The emotional tone of Q&As based on the informal lexicon of emotional phrases, referring to Yao Jiaquan et al. (2021)
Len	Number of characters in a Q&A	The natural logarithm of the number of characters in a research quiz
Vn	Number of questions and answers	Number of questions and answers in the current survey, one question and one answer count as 1
Year	Year	Annual dummy variables
Firm	Firm	Individual fixed effects

IndDire, *Dual*, *Foreign*, *Big4*, *Len* and *Vn* and a negative correlation significant at the 1 % level with *MB*, *Board*, *Cross* and *Tone*.

4.4. Regression analysis

To investigate the relationship between the medium of communication and investors' language concreteness, model (l) is used for regression and the results are reported in Table 6. In column (1), *MediaMode* and *Q_CON* show a significant positive correlation at the 1 % level without considering the control variables. After adding the control variables in column (2), the coefficient of *MediaMode* and *Q_CON* is 0.025 and is significant at the 5 % level, empirically supporting H1. In terms of economic significance, this result implies that the language of investors' questions is more concrete during online visits.

Among the control variables, larger companies (*Size*), higher gearing ratios (*Lev*) and higher growth (*Growth*) are associated with more concrete investor language (*Q_CON*). Conversely, a larger board of directors (*Board*) and a higher proportion of independent directors (*IndDire*) are linked to less concrete investor language (*Q_CON*). At the visit level, both *Vn* and *Len* are significantly positively correlated with *Q_CON* at the 1 % level. In contrast, *Tone* is significantly negatively correlated with *Q_CON*. This indicates that as the depth of discussions increases, investors' language concreteness increases, and they tend to be more specific in their questions when discussing more negative information.

In summary, both the statistical significance and economic significance of our results suggest that during online institutional visits, investors do not have the opportunity to engage in face-to-face communication to capture intangible information such as gestures and gazes. Therefore, they will try to obtain more information by increasing the concreteness of their questions, thereby compensating for the disadvantages of online communication in terms of information dissemination.

4.5. Robustness checks

To ensure the robustness of our conclusions and mitigate problems such as self-selection, the following tests are performed.

Table 3
Descriptive statistics.

Panel A: Descriptive statistics for all variables

Variable	Obs.	Mean	S.D.	Min	Median	Max
Q_CON	50,934	0.889	0.550	–1.099	0.981	1.946
MediaMode	50,934	0.077	0.266	0.000	0.000	1.000
Size	50,934	22.200	1.217	20.080	22.020	25.700
Lev	50,934	0.399	0.184	0.047	0.394	0.807
ROE	50,934	0.097	0.091	–0.268	0.092	0.358
Growth	50,934	0.232	0.345	–0.387	0.170	2.075
MB	50,934	0.309	0.135	0.071	0.290	0.718
Age	50,934	2.817	0.358	1.386	2.833	4.007
IndDire	50,934	0.376	0.055	0.333	0.333	0.571
Board	50,934	8.552	1.723	4.000	9.000	18.000
InsInve	50,934	0.403	0.247	0.007	0.416	0.873
SOE	50,934	0.227	0.419	0.000	0.000	1.000
Dual	50,934	0.338	0.473	0.000	0.000	1.000
Foreign	50,934	0.349	0.783	0.000	0.000	10.000
Big4	50,934	0.063	0.243	0.000	0.000	1.000
Cross	50,934	0.049	0.216	0.000	0.000	1.000
Tone	50,934	0.699	0.234	–0.947	0.753	0.993
Len	50,934	6.172	0.562	4.824	6.168	7.432
Vn	50,934	2.040	0.438	1.386	1.946	3.807

Panel B: Descriptive statistics by year for key variables

Year	**Variable**	**Obs.**	**Mean**	**S.D.**	**Min**	**Median**	**Max**
2012	*Q_CON*	1,981	0.924	0.526	–1.099	0.990	1.946
2013	*Q_CON*	6,202	0.900	0.532	–1.099	0.981	1.946
2014	*Q_CON*	6,705	0.895	0.531	–1.099	0.981	1.946
2015	*Q_CON*	5,510	0.864	0.554	–1.099	0.956	1.946
2016	*Q_CON*	5,978	0.882	0.559	–1.099	0.978	1.946
2017	*Q_CON*	5,230	0.885	0.549	–1.099	0.956	1.946
2018	*Q_CON*	4,614	0.901	0.552	–1.099	0.981	1.946
2019	*Q_CON*	4,326	0.894	0.561	–1.099	0.981	1.946
2020	*Q_CON*	4,609	0.896	0.564	–1.099	0.981	1.946
2021	*Q_CON*	5,779	0.875	0.567	–1.099	0.981	1.946
2012	*MediaMode*	1,981	0.008	0.087	0.000	0.000	1.000
2013	*MediaMode*	6,202	0.010	0.097	0.000	0.000	1.000
2014	*MediaMode*	6,705	0.011	0.106	0.000	0.000	1.000
2015	*MediaMode*	5,510	0.019	0.138	0.000	0.000	1.000
2016	*MediaMode*	5,978	0.021	0.144	0.000	0.000	1.000
2017	*MediaMode*	5,230	0.033	0.179	0.000	0.000	1.000
2018	*MediaMode*	4,614	0.045	0.208	0.000	0.000	1.000
2019	*MediaMode*	4,326	0.043	0.203	0.000	0.000	1.000
2020	*MediaMode*	4,609	0.324	0.468	0.000	0.000	1.000
2021	*MediaMode*	5,779	0.255	0.436	0.000	0.000	1.000

4.5.1. Heckman test

Investors may opt for online visits when analyzing companies with specific characteristics. To address potential endogeneity arising from self-selection bias, further tests are conducted using Heckman's two-stage approach. In the first stage, a probit model is used to estimate the factors that influence the choice of research method, applying a company's digital adoption score (*Dtech*) as an instrumental variable to calculate the inverse Mills ratio (*IMR*). The digital adoption score (*Dtech*) is a composite measure of a company's digital transformation, encompassing technological innovation, process innovation and business innovation. A higher *Dtech* score increases the likelihood of selecting online visits under comparable conditions. However, *Dtech* is not directly related to the concreteness of investors' Q&As, satisfying the requirements of an instrumental variable. In the second stage, the *IMR* obtained from the first stage is included as a control variable in the regression model to account for potential sample selection bias.

Regression (1) in Table 7 presents the first-stage results of the Heckman two-stage model. *Dtech* is positively and significantly associated with *MediaMode* at the 1 % level, confirming that firms with higher digital adoption are more inclined to conduct online visits. This validates the suitability of *Dtech* as an instrumental variable. Regression (2) reports the second-stage results. After controlling for sample selection bias by including *IMR*, the regression coefficient of *MediaMode* on *Q_CON* is 0.027 and remains significant at the 5 % level, further supporting H1.

4.5.2. PSM test

Currently, the penetration of online media in investor relations activities remains relatively low. Consequently, online visits constitute a small proportion of our sample, raising concerns about potential sample selection bias. To address this issue and mitigate endogeneity, we use the propensity score matching (PSM) method. Specifically, for each online visit, a corresponding site visit is identified based on the propensity score. The matched sample is then used to re-examine our research hypothesis.

First, the sample is divided based on *MediaMode*, with online visits (*MediaMode* = 1) designated as the treatment group and site visits (*MediaMode* = 0) as the control group. Nearest neighbor caliper matching is then applied, yielding 6,685 matched samples, as reported in Table 8. Panel A reports the results of the balance test. Compared with the pre-matching values, the standardized mean differences of all covariates are significantly reduced after matching. Additionally, the t-tests confirm that none of the matched variables exhibit

Table 4
Univariate analysis.

Variable	Site Visits		Online Visits		T-test
	Obs.	Mean	Obs.	Mean	
Q_CON	47,017	0.888	3,917	0.905	–0.017*
Size	47,017	22.168	3,917	22.564	–0.396***
Lev	47,017	0.399	3,917	0.402	–0.003
ROE	47,017	0.096	3,917	0.106	–0.010***
Growth	47,017	0.232	3,917	0.236	–0.004
MB	47,017	0.311	3,917	0.287	0.023***
Age	47,017	2.803	3,917	2.989	–0.186***
IndDire	47,017	0.376	3,917	0.380	–0.004***
Board	47,017	8.572	3,917	8.314	0.259***
InsInve	47,017	0.403	3,917	0.402	0.000
SOE	47,017	0.227	3,917	0.221	0.006
Dual	47,017	0.335	3,917	0.372	–0.037***
Foreign	47,017	0.333	3,917	0.538	–0.204***
Big4	47,017	0.062	3,917	0.075	–0.014***
Cross	47,017	0.051	3,917	0.026	0.025***
Tone	47,017	0.701	3,917	0.685	0.016***
Len	47,017	6.158	3,917	6.338	–0.180***
Vn	47,017	2.034	3,917	2.110	–0.075***

significant differences between the treatment and control groups, supporting the parallel trend assumption and validating the balance condition.

Furthermore, we perform a regression analysis based on the matched samples, with the results presented in Panel B. When *Q_CON* is used as the dependent variable, the regression coefficient of *MediaMode* is 0.030, significant at the 5 % level, providing additional support for H1.

4.5.3. Exclusion of a public health event

In 2020, the outbreak of COVID-19 profoundly affected economic activity and industrial operations, both domestically and globally. To curb the spread of the virus, numerous offline activities, including investor visits to listed companies, were suspended. As a result, some firms were forced to switch to online visits due to strict prevention and control policies, particularly between 18 January 2020 and 17 April 2020.

To eliminate the potential bias introduced by the pandemic, we exclude all samples from the outbreak period (18 January–17 April 2020). The regression results, reported in column (1) of Table 9, show that the coefficient of *MediaMode* on *Q_CON* is 0.021, significant at the 10 % level, confirming that our primary findings remain valid after removing the observations affected by the pandemic.

Additionally, as the impact of COVID-19 gradually diminished during the 2022–2023 period and economic activity resumed, we extend our sample period to 2012–2023 to further test the robustness of our results. Using the same method to construct our key variables, we obtain a total of 75,004 observations. The regression results, presented in column (2) of Table 9, indicate that the coefficient of *MediaMode* on *Q_CON* is 0.020, significant at the 1 % level, further reinforcing the robustness of our main findings.

4.5.4. Variable sensitivity test

In our previous regression analysis, the dependent variable constructed using the dictionary method focuses primarily on the frequencies of concrete and non-concrete words. Given that some concrete words are common and should be assigned lower weight, we perform TF-IDF processing and the values are standardized based on text length. Subsequently, the Q&A-level indicator is normalized to the [0, 1] range and, following the same procedure, a language concreteness indicator is constructed to replace our original dependent variable.

The regression results in Table 10 show that the coefficient of *MediaMode* on *Q_CON* is 0.001 and is significant at the 1 % level. Therefore, after replacing our explanatory variable, the results still support the conclusion that investors' questions are more concrete during online visits.

5. Heterogeneity analysis

Using the "5W" communication model, our study examines how different elements of institutional visits—such as the sender, message and receiver—affect investor engagement, with a specific focus on the concreteness of language in online and offline visits.

5.1. Sender: Privacy during institutional visits

Investors' online visits reduce the cost of information acquisition and enable simultaneous participation of many investors (Li and Zhao, 2021). However, increased participation implies less privacy (Wang et al., 2023). Social psychology research suggests that in larger groups, issues related to attribution and fairness may reduce individuals' sense of personal responsibility and commitment, leading to more generalized communication (Latané et al., 1979).

Using the number of participants as a measure of privacy during an investor's visit, *ResN* is defined as the negative logarithm of (the number of participants plus 1). A higher *ResN* value indicates a more private investor visit. The results in Table 11 show that the coefficient of the interaction term *MediaMode***ResN* is 0.013, which is significant at the 10 % level. This suggests that as privacy increases (i.e., there are fewer participants), investors' questions become more concrete, reinforcing the relationship between institutional visit settings and the linguistic precision of investors' questions.

Table 5
Correlation analysis.

	MediaMode	*Q_CON*	*Size*	*Lev*	*ROE*	*Growth*	*MB*	*Age*	*IndDire*	*Board*	*InsInve*	*SOE*	*Dual*	*Foreign*	*Big4*	*Cross*	*Tone*	*Len*	*Vn*
MediaMode	1.000																		
Q_CON	0.008*	1.000																	
Size	0.087***	–0.050***	1.000																
Lev	0.004	–0.004	0.554***	1.000															
ROE	0.029***	–0.022***	0.092***	–0.050***	1.000														
Growth	0.003	0.019***	–0.055***	0.054***	0.245***	1.000													
MB	–0.046***	–0.018***	0.021***	–0.355***	–0.181***	–0.113***	1.000												
Age	0.139***	–0.008*	0.263***	0.161***	0.031***	–0.034***	–0.044***	1.000											
IndDire	0.019***	–0.018***	0.066***	0.036***	–0.023***	0.010**	–0.015***	0.012***	1.000										
Board	–0.040***	–0.038***	0.264***	0.129***	0.075***	–0.036***	0.049***	0.092***	–0.482***	1.000									
InsInve	0.000	–0.002	0.378***	0.207***	0.164***	–0.025***	–0.118***	0.125***	–0.083***	0.240***	1.000								
SOE	–0.004	–0.014***	0.348***	0.182***	0.009**	–0.075***	0.034***	0.220***	–0.093***	0.324***	0.404***	1.000							
Dual	0.021***	–0.013***	–0.068***	–0.034***	0.002	0.008*	–0.041***	–0.097***	0.127***	–0.144***	–0.147***	–0.242***	1.000						
Foreign	0.069***	–0.007	0.147***	0.057***	0.026***	0.008*	–0.058***	0.107***	0.050***	0.089***	0.079***	0.009**	0.031***	1.000					
Big4	0.015***	–0.055***	0.428***	0.192***	0.064***	–0.022***	–0.012***	0.064***	0.062***	0.221***	0.187***	0.140***	–0.005	0.207***	1.000				
Cross	–0.030***	–0.064***	0.406***	0.194***	0.017***	–0.026***	0.071***	0.121***	0.118***	0.254***	0.144***	0.191***	0.008*	0.152***	0.544***	1.000			
Tone	–0.018***	–0.088***	0.012***	0.003	0.060***	0.068***	–0.059***	0.040***	0.022***	0.010**	–0.043***	–0.021***	0.020***	0.024***	0.018***	0.010**	1.000		
Len	0.086***	0.049***	0.095***	0.061***	–0.030***	0.044***	–0.006	0.131***	0.058***	–0.025***	–0.076***	–0.029***	0.001	0.070***	0.029***	–0.008*	0.233***	1.000	
Vn	0.046***	0.035***	–0.145***	–0.143***	0.066***	0.045***	–0.013***	–0.055***	–0.007	–0.039***	–0.033***	–0.083***	–0.00200	–0.011**	–0.078***	–0.082***	–0.108***	–0.303***	1.000

Table 6
Communication media and investor language concreteness.

	Q_CON (1)	(2)
MediaMode	0.042***	0.025**
	(3.71)	(2.25)
Size		0.026***
		(2.67)
Lev		0.085**
		(2.12)
ROE		–0.007
		(–0.18)
Growth		0.015*
		(1.74)
MB		–0.053
		(–1.44)
Age		0.214***
		(4.70)
IndDire		–0.509***
		(–4.88)
Board		–0.012***
		(–2.60)
InsInve		–0.015
		(–0.47)
SOE		0.018
		(0.67)
Dual		0.008
		(0.75)
Foreign		0.006
		(1.40)
Big4		–0.001
		(–0.05)
Cross		–0.163
		(–0.86)
Tone		–0.217***
		(–18.66)
Len		0.063***
		(9.57)
Vn		0.059***
		(7.65)
_Cons	0.886***	–0.382
	(368.58)	(–1.55)
N	50,798	50,798
Year Fixed	Yes	Yes
Firm Fixed	Yes	Yes
R^2	0.175	0.184
Adj R^2	0.145	0.154

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

5.2. Message: Information uniqueness and board secretary traits

5.2.1. Information uniqueness

From the perspective of investor demand for information, professional research reports provided by analysts serve as a key channel for investors to obtain valuable information. As information intermediaries, ana-

Table 7
Heckman test.

	(1) *MediaMode*	(2) *Q_CON*
Dtech	0.005***	
	(9.34)	
MediaMode		0.027**
		(2.37)
IMR		–0.143***
		(–3.69)
Controls	Yes	Yes
_Cons	–6.711***	0.386
	(–18.21)	(1.19)
N	50,860	50,723
Year Fixed	Yes	Yes
Firm Fixed	Yes	Yes
R^2		0.185
Adj R^2		0.154

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

lysts' reports influence investors' decision-making (Li and Zhang, 2019). When investors can access the necessary information through alternative sources, the uniqueness of information discovered during institutional visits is likely to diminish.

To test this conjecture, we use the negative of the number of analyst reports issued for a company in a given year (*InfUnig1*) and the negative of the number of analysts following the company (*InfoUnig2*). A higher value for these two variables indicates that the information exchanged during institutional visits is more unique. The regression results are shown in Table 12. The coefficient of *MediaMode* InfUnig1* is 0.018, which is significant at the 10 % level, while the coefficient of *MediaMode* InfoUnig2* on *Q_CON* is 0.028, significant at the 5 % level. These findings confirm that when investors can obtain more professional information or interpretations from analysts, the uniqueness of the information exchanged during investors' online visits decreases. As a result, investors' questions are less concrete when their information needs are largely met through alternative channels.

5.2.2. Board secretary traits

The board secretary plays a key role in corporate information disclosure and communication with investors, directly shaping the content and the manner in which market participants access company information. Prior research shows that board secretaries can influence investor sentiment by enhancing the positive tone of financial reports, thereby boosting investor motivation (Liu et al., 2022). In institutional investor visits, the board secretary typically represents the listed company, so all investor engagement transcripts are compiled and disclosed under their supervision.

To examine how board secretary traits moderate the relationship between the medium of communication and investor language concreteness, we introduce two variables: a dummy variable indicating whether the board secretary holds shares in the company (*SecreShare*) and the number of shares held by the secretary (*Sharend*). The regression results in Table 13 show that the coefficient of *MediaMode*SecreShare* is –0.037 and the coefficient of *MediaMode*Sharend* is –0.003, both significant at the 10 % level.

These findings suggest that the board secretary's shareholding weakens the positive effect of investors' online visits on their linguistic concreteness. Furthermore, this mitigating effect becomes more pronounced as the proportion of shares held by the secretary increases. These results highlight the essential role of the board secretary in the disclosure of information. When the secretary holds shares in the company, investors

Table 8
Propensity score matching (PSM).

Panel A: Analysis of PSM parallelism

Variable	Unmatched	Mean		Standardized mean difference	Change in standardized mean difference (%)	T-test		V(T) / V(C)
	Matched	Experimental group	Control group			t	p > |t|	
Size	U	22.564	22.168	33.900		19.620	0.000	0.83*
	M	22.563	22.602	–3.300	90.100	–1.410	0.158	0.70*
Lev	U	0.402	0.399	1.700		0.970	0.331	0.73*
	M	0.402	0.403	–0.800	54.600	0.340	0.734	0.69*
ROE	U	0.106	0.096	9.900		6.550	0.000	1.50*
	M	0.106	0.108	–2.500	75.000	–1.060	0.289	1.26*
Growth	U	0.236	0.232	1.200		0.740	0.461	0.94
	M	0.236	0.235	0.200	81.000	–0.100	0.918	0.89*
MB	U	0.287	0.311	–17.60		–10.380	0.000	0.92*
	M	0.287	0.290	–1.700	90.400	–0.770	0.443	1.01
Age	U	2.989	2.803	56.100		31.570	0.000	0.71*
	M	2.989	2.994	–1.500	97.300	–0.770	0.440	1.14*
IndDire	U	0.380	0.376	7.200		4.350	0.000	1.06
	M	0.380	0.380	–1.000	85.400	–0.460	0.646	1.02
Board	U	8.314	8.572	–16.000		–9.030	0.000	0.72*
	M	8.314	8.319	–0.300	98.000	0.150	0.884	0.81*
InsInve	U	0.402	0.403	–0.200		–0.110	0.915	1.03
	M	0.402	0.406	–1.700	–833.400	–0.730	0.463	1.04
SOE	U	0.221	0.227	–1.500		–0.920	0.358	.
	M	0.221	0.229	–1.900	–24.700	–0.850	0.397	.
Dual	U	0.372	0.335	7.700		4.670	0.000	.
	M	0.372	0.373	–0.100	98.100	–0.060	0.950	.
Foreign	U	0.538	0.333	23.000		15.720	0.000	1.72*
	M	0.537	0.517	2.200	90.300	0.870	0.385	0.94*
Big4	U	0.075	0.062	5.400		3.370	0.001	.
	M	0.075	0.076	–0.300	94.300	–0.130	0.896	.
Cross	U	0.026	0.051	–12.800		–6.880	0.000	.
	M	0.026	0.027	–0.500	96.400	–0.250	0.806	.
Tone	U	0.685	0.701	–6.800		–4.000	0.000	0.91*
	M	0.685	0.691	–2.500	63.400	–1.100	0.272	0.92*
Len	U	6.338	6.158	33.000		19.370	0.000	0.89*
	M	6.338	6.332	1.300	96.200	0.560	0.576	0.91*
Vn	U	2.110	2.034	17.100		10.380	0.000	1.05
	M	2.110	2.105	1.000	94.100	0.430	0.666	0.95

Panel B: Regression after PSM

	Q_CON
MediaMode	0.030**
	(2.10)
Controls	Yes
_Cons	–0.373
	(–0.67)
N	14,433
Year Fixed	Yes
Firm Fixed	Yes
R^2	0.245
Adj R^2	0.167

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

Table 9
Excluding the objective impact of a public health event.

	Q_CON (1)	(2)
MediaMode	0.021*	0.020***
	(1.76)	(2.86)
Controls	Yes	Yes
_Cons	–0.420*	0.154
	(–1.70)	(0.88)
N	50,090	74,681
Year Fixed	Yes	Yes
Firm Fixed	Yes	Yes
R^2	0.186	0.227
Adj R^2	0.155	0.197

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

Table 10
Alternative measure of the dependent variable.

	Q_CON
MediaMode	0.001***
	(4.73)
Controls	Yes
_Cons	–0.593***
	(–120.10)
N	50,798
Year Fixed	Yes
Firm Fixed	Yes
R^2	0.181
Adj R^2	0.151

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

may perceive corporate disclosures as more credible and interpret the information more efficiently, reducing their reliance on concrete language when formulating questions during online visits.

5.3. Receiver: Management traits

As strategic communicators in interactions with investors, management engagement directly influences the dynamics of information flow. Building on managers' dual role as information gatekeepers and corporate narrators, we construct a measure of management language concreteness (*A_CON*) using the same method applied to construct investor language concreteness (*Q_CON*). A higher value of *A_CON* indicates a higher level of concreteness in management responses during investor visits.

To examine how the medium of communication affects the concreteness of management language, we conduct a regression using *A_CON* as the dependent variable. The results in column (1) of Table 14 show that *MediaMode* is negatively and significantly correlated with *A_CON* at the 1 % level, indicating that management's language is less concrete during investors' online visits than during site visits. This effect may be driven by the constraints of online communication, which encourage management to adopt more generalized language. Meanwhile, the transparency and recording capabilities of online channels encourage management to provide more cautious and vague responses to mitigate the risks of unintended information leaks. This finding highlights the strategic interplay between investors and management in information exchange, where both parties adjust their communication tactics based on the medium.

Prior research suggests that management strategically adjusts its responses to investor questions on public platforms based on the tone of the questions (Wu et al., 2022), with empirical evidence demonstrating positive market reactions to management's use of concrete language (Elliott et al., 2015; Pan et al., 2018). Building on this bidirectional influence framework, we hypothesize that management language concreteness (*A_CON*) inversely moderates investor language concreteness (*Q_CON*). As shown in column (2) of Table 14, the sta-

Table 11
Communication media, institutional visit privacy and investor language concreteness.

	Q_CON
MediaMode	0.057**
	(2.54)
ResN	–0.004
	(–1.20)
*MediaMode*ResN*	0.013*
	(1.86)
Controls	Yes
_cons	–0.355
	(–1.44)
N	50,547
Year Fixed	Yes
Firm Fixed	Yes
R^2	0.185
Adj R^2	0.154

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

Table 12
Communication media, information uniqueness and investor language concreteness.

	Q_CON	
	(1)	(2)
MediaMode	0.076**	0.089***
	(2.40)	(2.82)
InfoUniq1	–0.018***	
	(–4.18)	
MediaMode InfoUniq1*	0.018*	
	(1.92)	
InfoUniq2		–0.023***
		(–4.20)
MediaMode InfoUniq2*		0.028**
		(2.39)
Controls	Yes	Yes
_cons	–0.364	–0.353
	(–1.34)	(–1.30)
N	46,950	46,950
Year Fixed	Yes	Yes
Firm Fixed	Yes	Yes
R^2	0.189	0.189
Adj R^2	0.159	0.159

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

Table 13
Communication media, board secretary traits and investor language concreteness.

	Q_CON (1)	(2)
MediaMode	0.043***	0.043***
	(2.59)	(2.60)
SecreShare	–0.012	
	(–1.41)	
*MediaMode*SecreShare*	–0.037*	
	(–1.75)	
Sharend		–0.001
		(–1.52)
*MediaMode*Sharend*		–0.003*
		(–1.77)
Controls	Yes	Yes
_cons	–0.349	–0.352
	(–1.39)	(–1.40)
N	49,181	49,181
Year Fixed	Yes	Yes
Firm Fixed	Yes	Yes
R^2	0.187	0.187
Adj R^2	0.156	0.156

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

Table 14
Communication media, management traits and investor language concreteness.

	A_CON (1)	*Q_CON* (2)	*Q_CON* (3)
MediaMode	–0.071***	0.088***	0.015
	(–6.73)	(2.71)	(1.21)
A_CON		0.160***	
		(30.28)	
*MediaMode*A_CON*		–0.022*	
		(–1.71)	
Chairman			–0.040***
			(–4.65)
*MediaMode*Chairman*			0.056**
			(2.18)
Controls	Yes	Yes	Yes
_cons	–2.733***	0.047	–0.390
	(–12.04)	(0.20)	(–1.59)
N	50,798	50,798	50,798
Year Fixed	Yes	Yes	Yes
Firm Fixed	Yes	Yes	Yes
R^2	0.674	0.203	0.185
Adj R^2	0.662	0.173	0.154

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

Table 15
Communication media, investor language concreteness and analyst forecasts.

	FERROR		*FOPTI*	
	(1)	(2)	(3)	(4)
Presence of analysts	YES	NO	YES	NO
Q_CON	0.020**	–0.028**	0.021**	–0.033**
	(2.03)	(–2.00)	(2.06)	(–2.19)
MediaMode	0.045*	–0.016	0.043*	–0.046
	(1.81)	(–0.59)	(1.72)	(–1.60)
MediaMode Q_CON*	–0.049**	0.006	–0.043*	0.036
	(–2.16)	(0.28)	(–1.81)	(1.45)
Controls	Yes	Yes	Yes	Yes
_cons	–0.523	0.555	–0.564	0.273
	(–0.86)	(0.59)	(–0.89)	(0.28)
N	8,822	2,222	8,822	2,222
Year Fixed	Yes	Yes	Yes	Yes
Firm Fixed	Yes	Yes	Yes	Yes
Difference between groups		0.035		0.023
R^2	0.221	0.355	0.229	0.359
Adj R^2	0.121	0.227	0.130	0.231

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively. The t-statistics are shown in brackets.

tistically significant negative coefficient of *MediaMode*A_CON* on *Q_CON* indicates that when management provides more concrete answers during online visits, investors' questions become less concrete.

In addition, the literature suggests that greater management attention during interactions with investors may impact the accuracy of analysts' earnings forecasts (Luo and Li, 2022). To assess the effect of management attention on information dissemination, we use the presence of the chairman (*Chairman*) as a proxy for management involvement and examines its moderating effect. The regression results in column (3) of Table 14 show that the coefficient of *MediaMode*Chairman* on *Q_CON* is 0.056, which is significant at the 5 % level. This suggests that when top management is present during interactions with investors, investors tend to ask more concrete questions, likely due to increased expectations of valuable information.

Overall, these findings underscore the critical role of management traits in shaping the dynamics of investor visits. The choice of communication medium, management language concreteness and the presence of senior executives all influence the depth and concreteness of investors' questions, reflecting a strategic exchange of information between firms and market participants.

6. Further analysis

As key information intermediaries in capital markets and primary users of corporate disclosures (Shi et al., 2021), analysts leverage their expertise and information advantages to assess corporate performance and other relevant factors. By publishing forecast reports, they help to mitigate information asymmetry between investors and firms, thereby enhancing capital market efficiency. Research on conference calls shows that analysts effectively revise their forecasts based on positive and negative signals embedded in management's linguistic complexity (Brian and Huang, 2024). Analysts primarily influence the market through earnings forecast reports, and prior studies suggest that site visits help them overcome language barriers, improving forecast accuracy (Tam and Tian, 2023). Conversely, online visits may lead to lower EPS forecast accuracy due to the lack of direct on-site observation (Kang and Li, 2024). This raises the question of whether the use of specific language can compensate for the information limitations of virtual settings.

Following Cheng et al. (2016), we construct an explanatory variable for analyst forecast bias (*FERROR*) to measure changes in forecast deviations before and after investor visits. Specifically, *FERROR* is defined as the difference between the mean forecast bias of analysts' most recent predictions 15 days after a visit and those 15 days before a visit: $FERROR = FERROR\_After_{i,t} - FERROR\_Before_{i,t}$. A higher value of *FERROR* indicates greater forecast bias after visits, implying reduced forecast accuracy. Here, $FERROR\_After_{i,t}$ represents the mean deviation of all analyst forecasts in the 15 days following visit t to firm i, while $FERROR\_Before_{i,t}$ is the mean deviation of all analyst forecasts in the 15-day window preceding the visit. The deviation measure is calculated as $ferror = \frac{|FEPS-AEPS|}{AEPS}$, where *FEPS* is the analysts' forecast EPS for firm i in a given year and *AEPS* is the actual EPS for the same period. Additionally, we use the analyst forecast optimism change indicator (*FOPTI*), where optimism is defined as $optim = \frac{FEPS-AEPS}{AEPS}$. This metric accounts for both the direction and magnitude of forecast bias.

Table 15 presents the regression results regarding the impact of specific investor questions during online visits on analyst behavior, distinguishing subgroups based on analyst participation. In the subgroup with analyst participation, the coefficient on *MediaMode* is positive and significant at the 10 % level for *FERROR*, indicating that analyst involvement in online visits tends to increase forecast bias, consistent with prior findings (Kang and Li, 2024). However, the interaction term *MediaMode**Q_CON* is negative and significant at the 5 % level, suggesting that greater linguistic concreteness in investors' questions mitigates this forecast bias. In contrast, the interaction term is positive but not statistically significant for the subgroup without analyst participation.

Similarly, in the regression for *FOPTI*, the interaction term *MediaMode**Q_CON* exhibits a significant and negative correlation at the 10 % level when analysts are present, whereas it remains positive but not significant in the subgroup without analysts. This implies that analysts, as specialized market participants, are able to extract and integrate specific information from investor visits, leading to more precise and rational earnings forecasts.

7. Conclusion

The rise of digital technologies has fundamentally transformed capital market communications, shifting institutional investor visits from traditional site visits to a blended model incorporating online interactions. This dual-channel system reflects both technological progress and market adaptability. These complementary communication modes coexist as key drivers of financial information exchange in the capital market.

Through textual analysis of investor engagement transcripts, we measure linguistic concreteness to study how market participants adapt to digital communication constraints. Our findings reveal that investors demonstrate increased communication precision in online settings, using more concrete language than during site visits. In addition, investor language concreteness varies with the levels of privacy of visits, the uniqueness of information, the characteristics of board secretaries and management responsiveness. Finally, analyst forecast accuracy increases with linguistic precision in digital communications, showing how the strategic use of language can overcome media limitations.

Our findings extend Lasswell's communication framework to the corporate disclosure context, which not only broadens the research boundaries on the impact of digital technology on communication media but also helps us understand the balancing mechanism between information transmission willingness and media characteristics. Based on the results of our study, we propose the following recommendations.

First, regulators should actively support diverse methods of interaction. The growing prevalence of online media in capital markets is an inevitable trend, given the advantages they offer in terms of cost-effectiveness, accessibility and broad participation. Although online media may lack some nonverbal cues, our findings show that investors effectively compensate with concrete language. Therefore, policymakers should take an open stance toward online communication, supporting its application and encouraging capital markets to explore diverse methods of interaction.

Second, enhancing the information disclosure framework for online media is crucial. The widespread use of online communication has introduced new challenges to information transparency, requiring stricter oversight. Regulatory bodies should establish clear guidelines for online visits to ensure compliance and transparency. It is recommended to standardize online communication channels to facilitate efficient interactions between investors and listed companies. Additionally, leveraging artificial intelligence and big data technologies can enhance supervisory capabilities, improving multi-channel monitoring of information disclosure.

Third, listed companies must proactively adapt to digital transformation. Firms should actively use online media, promoting greater participation and interactivity in online settings. By refining the concreteness and precision of their responses, companies can maximize the role of online visits in information dissemination and corporate oversight. Strengthening online engagement not only enhances corporate governance transparency but also boosts market confidence and increases firm value.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. **Appendix 1.** Examples of concrete vocabulary

Number	Part of speech					
	Verb	Number	Past focused	Adj	Non-specific quantifiers	Future focused
01	Form	0	Original	Mature	Limited	Follow-up
02	Expand	1	Near term	Smooth	Some	Second half
03	Produce	1	Completion	Thereby	Approximately	Prospects
04	Begin	2	Third quarter	Tremendous	Strengthen	Accelerate
05	Become	3	Past	Direct	Basic	Future

Appendix 2. Statistics on the annual mean of the six categories of concrete words used by investors.

	Investor questions					
Year	Verb	Number	Past focused	Adj	Non-specific quantifiers	Future focused
2012	4.814	0.042	0.814	2.669	0.432	0.841
2013	4.829	0.040	0.816	2.716	0.432	0.863
2014	4.961	0.034	0.795	2.744	0.441	0.936
2015	4.884	0.037	0.766	2.672	0.424	0.969
2016	4.993	0.036	0.811	2.721	0.457	0.989
2017	5.062	0.039	0.854	2.819	0.467	0.996
2018	5.271	0.047	0.884	2.967	0.490	1.023
2019	5.430	0.042	0.871	3.021	0.518	1.087
2020	5.720	0.046	0.957	3.300	0.565	1.165
2021	5.728	0.047	0.906	3.283	0.570	1.167

Appendix 3. Statistics on the annual mean of the six categories of concrete words used by management.

	Management responses					
Year	Verb	Number	Past focused	Adj	Non-specific quantifiers	Future focused
2012	23.970	0.563	3.320	13.780	2.605	4.759
2013	26.480	0.618	3.501	15.140	2.881	5.431
2014	29.210	0.588	3.654	16.170	2.990	6.130
2015	34.910	0.739	4.092	18.630	3.270	7.390
2016	38.420	0.842	4.473	20.460	3.717	8.106
2017	39.760	0.939	4.753	21.680	4.042	8.534
2018	39.790	0.937	4.719	21.960	4.062	8.558
2019	41.870	0.932	4.771	22.650	4.206	9.189
2020	44.540	0.955	5.561	24.990	4.626	9.471
2021	43.840	0.922	5.247	24.300	4.577	9.608

References

Bowen, R.M., Dutta, S., Tang, S., Zhu, P., 2018. Inside the "black box" of private in-house meetings. Rev. Acc. Stud. 23, 487–527.

Bowen, R.M., Dutta, S., Tang, S., Zhu, P., 2023. Does corporate governance quality influence insider trading around private meetings between managers and investors? Account. Horiz. 37 (3), 27–57.

Bozeman, D.P., Kacmar, K.M., 1997. A cybernetic model of impression management processes in organizations. Organ. Behav. Hum. Decis. Process. 69 (1), 9–30.

Bu, J., Sun, G.G., 2020. Investor site visits corporate violations in listed companies: mechanisms and empirical tests. Account. Res. 5, 30–47 (in Chinese).

Cai, X., Jiang, F., Kang, J.K., 2023. Remote board meetings and board monitoring effectiveness: evidence from China. Rev. Financ. Stud. 36 (11), 4318–4372.

Cao, J., Wang, H., Zhou, S., 2022. Soft activism and corporate dividend policy: evidence from institutional investors site visits. Finance 75, 102221.

Carlson, J.R., Robert, W.Z., 1999. Channel expansion theory and the experiential nature of media richness perceptions. Acad. Manag. J. 42 (2), 153–170.

Chen, Q.Y., Xue, F., Yang, S.Z., et al., 2021. "Seeking opportunities in crisis": can online communication improve information efficiency in the capital market? Econ. Manag. 43 (5), 142–158 (in Chinese).

Cheng, Q., Du, F., Wang, X., et al., 2016. Seeing is believing: analysts' corporate site visits. Rev. Acc. Stud. 21, 1245–1286.

Cheng, X.K., Li, H.Y., Gao, S.H., 2017. Institutional investor visits and management's earnings forecast methods. Manag. Sci. 30 (1), 131–145 (in Chinese).

Cheng, Q., Du, F., Wang, B.Y., et al., 2019. Do corporate site visits impact stock prices? Contemp. Account. Res. 36 (1), 359–388.

Daft, R.L., Lengel, R.H., 1986. Organizational information requirements, media richness and structural design. Manag. Sci. 32 (5), 554–571.

Elliott, W.B., Rennekamp, K.M., White, B.J., 2015. Does concrete language in disclosures increase willingness to invest? Rev. Acc. Stud. 20, 839–865.

Estes, R.W., 1982. The auditor's report and investor behavior. Lexington Books, Lexington. MA.

Gao, H., Huang, J., Zhang, T., 2020. Can online annual general meetings increase shareholders' participation in corporate governance? Financ. Manag. 49 (4), 1029–1050.

Han, B., Kong, D., Liu, S., 2018. Do analysts gain an informational advantage by visiting listed companies? Contemp. Account. Res. 35 (4), 1843–1867.

Hansen, J., Wänke, M., 2010. Truth from language and truth from fit: the impact of linguistic concreteness and level of construal on subjective truth. Pers. Soc. Psychol. Bull. 36 (11), 1576–1588.

Hiltz, S.R., Johnson, K., Turoff, M., 1986. Experiments in group decision making communication process and outcome in face-to-face versus computerized conferences. Hum. Commun. Res. 13 (2), 225–252.

Huang, W.Y., 2018. A review of international research on early dual coding theory. Southeast Commun. 2, 40–43 (in Chinese).

Kang, X., Li, Y., 2024. Seeing is believing: are investors' corporate online visits as informative as site visits? Appl. Econ. Lett., 1–6

Kimura, S., Ooseki, E., Aburakawa, Y., et al., 2021. Evaluation and formulation of the sense of social telepresence in video-mediated communication systems: contribution of eye contact to enhancing social telepresence. J. Soc. Inf. Disp. 29 (3), 179–195.

Kruger, J.M., Chowers, I., 2020. The ethical advantages of video conferencing in medical education. Med. Educ. Online 25 (1), 1787310.

Langacker, R.W., 1987. Nouns and verbs. Language, 53–94.

Larrimore, L., Jiang, L., Larrimore, J., et al., 2011. Peer to peer lending: the relationship between language features, trustworthiness, and persuasion success. J. Appl. Commun. Res. 39 (1), 19–37.
Latané, B., Williams, K., Harkins, S., 1979. Many hands make light the work: the causes and consequences of social loafing. J. Pers. Soc. Psychol. 37 (6), 822.
Li, Y.Z., Zhang, L., 2019. Accounting Information Quality and Investment Decision: From the Perspective of Heterogeneous Beliefs. Accoun. Mon. 24, 74–82.
Li, Y., Zhao, L., 2021. The value of online research: from the perspective of total factor productivity of enterprises. Financ. Res. 47 (12), 136–149 (in Chinese).
Li, H., Yang, M., Chan, K.C., et al., 2022. Do institutional investors' corporate site visits impact seasoned equity offering discounts? Evidence from detailed investor bids in SEO auctions. Res. Int. Bus. Financ. 62, 101746.
Liu, S., Dai, Y., Kong, D., 2017. Does it pay to communicate with firms? Evidence from firm site visits of mutual funds. J. Bus. Financ. Acc. 44 (5–6), 611–645.
Liu, J.W., Deng, B.F., Ji, L., 2022. Is the information communication efficiency higher after the board secretary holds shares? Based on the empirical evidence of Chinese A-share listed companies. Foreign Econ. Manag. 44 (2), 36–51.
Lu, D., Yu, D., Yang, D., 2019. Financial report readability, investor field visits, and hedging strategies. Account. Res. 10, 34–41 (in Chinese).
McClelland, J.L., Rumelhart, D.E., 2020. Distributed memory and the representation of general and specific information. Connect. Psychol. Psychology Press, 75–106.
Nilsson, M., Mattes, J., 2015. The spatiality of trust: Factors influencing the creation of trust and the role of face-to-face contacts. Eur. Manag. J. 33 (4), 230–244.
Paivio, A., 1991. Dual coding theory: retrospect and current status. Can. J. Psychol. 45 (3), 255.
Pan, L., McNamara, G., Lee, J.J., et al., 2018. Give it to us straight (most of the time): top managers' use of concrete language and its effect on investor reactions. Strateg. Manag. J. 39 (8), 2204–2225.
Parker, E.B., Short, J., Williams, E., et al., 1976. The social psychology of telecommunications. Contemp. Sociol. 7 (1), 32.
Peng, Q.P., Zhong, X., Zhou, H.K., 2022. Double-edged sword effect of institutional investor visits on corporate innovation. J. Manag. Eng. 36 (4), 108–117 (in Chinese).
Purvanova, R.K., 2014. Face-to-face versus virtual teams: what have we really learned? Psychologist-Manager J. 17 (1), 2.
Sadoski, M., 2001. Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. Educ. Psychol. Rev. 13, 263–281.
Semin, G.R., Fiedler, K., 1988. The cognitive functions of linguistic categories in describing persons: social cognition and language. J. Pers. Soc. Psychol. 54 (4), 558.
Shi, X.Y., Bu, D.L., Wang, Y.H., 2021. Does "Shifting from Real to Virtual" Affect the Information Environment of Capital Markets? Based on the Perspective of Analyst Tracking. Contemp. Financ. Econ. 438 (5), 127–137.
Snefjella, B., Kuperman, V., 2015. Concreteness and psychological distance in natural language use. Psychol. Sci. 26 (9), 1449–1460.
Solomon, D., Soltes, E., 2015. What are we meeting for? The consequences of private meetings with investors. J. Law Econ. 58 (2), 325–355.
Song, Y., Xian, R., 2024. Institutional investors' corporate site visits and firm-level climate change risk disclosure. Int. Rev. Financ. Anal. 93, 103145.
Storper, M., Venables, A.J., 2004. Buzz: face-to-face contact and the urban economy. J. Econ. Geogr. 4 (4), 351–370.
Szeto, E., Cheng, A.Y.N., 2016. Towards a framework of interactions in a blended synchronous learning environment: what effects are there on students' social presence experience? Interact. Learn. Environ. 24 (3), 487–503.
Tan, S.T., Cui, X.Y., 2015. Can visits to listed companies improve analyst forecast accuracy? World Econ. 38 (4), 126–145 (in Chinese).
Tan, J.S., Lin, Y.C., 2016. Governance effects of institutional investors on information disclosure: evidence from institutional visit behavior. Nankai Bus. Rev. 19 (5), 115–126 (in Chinese).
Tam, L.H.K., Tian, S., 2023. Language barriers, corporate site visit, and analyst forecast accuracy. Q. Rev. Econ. Finance 91, 68–83.
Toma, C.L., D'Angelo, J.D., 2015. Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. J. Lang. Soc. Psychol. 34 (1), 25–45.
Wang, J.J., Liu, D., Liu, F.F., 2023. Joint visits and new values of "relationship-oriented" analysts. Financ. Res. 49 (3), 94–109 (in Chinese).
Wang, Q., Cao, Z., Lee, E., 2024. Do corporate site visits affect the informational role of independent directors? Financ. Manag. 53 (4), 867–903.
Williams, E., 1977. Experimental comparisons of face-to-face and mediated communication: a review. Psychol. Bull. 84 (5), 963.
Wu, D., Gao, S., Chan, K.C., et al., 2022. Do Firms Strategically Respond to Retail Investors on the Online Interactive Information Disclosure Platform? Financ. Res. Lett. 47, 102631.
Xiang, C., Yang, J., 2022. Impact of field visit information disclosure on informed trading probability of stocks. Manag. Sci. 35 (4), 127–142 (in Chinese).
Xiao, B.Q., Peng, Y., Fang, L.B., et al., 2017. Are visits to listed companies useful for investment decisions? Empirical research based on analyst visit reports. Nankai Bus. Rev. 20 (1), 119–131 (in Chinese).
Xu, Z.L., Gao, L., Lin, Y.C., 2021. Buy-side institutional visits and stock excess returns. Syst. Eng. Theory Pract. 41 (10), 2457–2475.
Yao, J.Y., Feng, X., Wang, Z.J., et al., 2021. Tone, emotion, and market impact: based on a financial sentiment dictionary. J. Manag. Sci. 24 (5), 26–46 (in Chinese).

Yan, H.M., Zhu, Z.L., Xiong, H.L., 2024. Institutional investor visits and company audit opinion shopping behavior: stimulating effect or inhibiting effect? J. Manag. Eng. 38 (2), 104–120 (in Chinese).
Yue, S., Dong, D., Wu, F., et al., 2023. More pay more gain? —Empirical research on fund management corporation visiting listed company and its fund performance. Int. J. Financ. Econ. 28 (1), 169–176.
Zhang, G.L., 2004. Principles of Communication. Fudan University Press (in Chinese).
Zhang, Y., Lyu, Q., Han, Y., 2024. Institutional investors drive corporate green governance: supervisory effects and internal mechanisms. Manag. World 40 (4), 197–221 (in Chinese).
Zhao, W., Yang, H., Zhou, H., 2022. Linguistic specificity and stock price synchronicity. China J. Account. Res. 15 (1), 100219.

China Journal of Accounting Research 18 (2025) 100432

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Fruitless effort? The effects of risk management disclosure specificity on nonprofessional investors' judgments☆

Xinyan Miao [a], Shaofei Wang [b,c,d], Yue Wang [d], Bo Zhou [d,e,*]

[a] *School of Accountancy, Shanghai University of Finance and Economics, China*
[b] *College of Business, Shanghai University of Finance and Economics, China*
[c] *School of Advanced Accounting and Auditing, Shanghai University of Finance and Economics, China*
[d] *Institute of Accounting and Finance, Shanghai University of Finance and Economics, China*
[e] *Dishui Lake Advanced Finance Institute, Shanghai University of Finance and Economics, China*

ARTICLE INFO

Article history:
Received 2 August 2024
Accepted 14 May 2025
Available online 05 June 2025

JEL Classification:
C91
D81
D83
G11
G41
M41

ABSTRACT

The widespread use of boilerplate disclosures in financial reporting has led regulators to express concerns about the decision-usefulness of annual reports. We conduct two experiments to determine the effects of risk management disclosure presence and specificity on nonprofessional investors' judgments. In Experiment 1, we manipulate risk management disclosure at three levels (non-disclosure vs. generic or specific risk management disclosures). Relative to the non-disclosure condition, nonprofessional investors exhibit more favorable investment judgments when provided with a specific risk management disclosure. However, generic disclosure has a negligible influence on nonprofessional investors' judgments. We find no convincing evidence supporting potential alternative explanations. Experiment 2 confirms this mechanism and provides further evidence that the observed effects are not driven by the lengths of specific disclosures.

* Corresponding author at: Institute of Accounting and Finance, Shanghai University of Finance and Economics, 777 Guoding Road, Yangpu District, Shanghai, People's Republic of China.

E-mail addresses: musesings@163.sufe.edu.cn (X. Miao), wang.shaofei@mail.shufe.edu.cn (S. Wang), wang.yue@mail.shufe.edu.cn (Y. Wang), zhou.bo@mail.shufe.edu.cn (B. Zhou).

☆ The study was approved by the ethics committee of the University of the authors, all participants gave written informed consent, and their rights are protected.

https://doi.org/10.1016/j.cjar.2025.100432

Keywords:
Risk management disclosure specificity
Non-financial disclosure
Investment judgments
Scenario thinking

1. Introduction

Increasingly, capital markets are emphasizing the importance of risk information disclosure, particularly its relevance for investors and other stakeholders (Campbell et al., 2014; Elshandidy & Neri, 2015; Ibrahim & Aboud, 2023). However, in practice, firms' risk disclosures differ in various aspects, such as tone (Wang, 2021; Elshandidy & Zeng, 2022; Tan & Yeo, 2023), readability (Linsley & Lawrence, 2007; Jia & Li, 2022; Yao et al., 2024) and specificity (Kravet & Muslu, 2013; Campbell et al., 2014; Cazier et al., 2021; Arikan, 2021). Many regulatory bodies require specificity in risk information disclosures. For example, the China Securities Regulatory Commission (CSRC, 2012) requires companies to disclose risk information using specific language, rather than generic boilerplate, aiming to provide investors with useful information. Likewise, the United States (U.S.) Securities and Exchange Commission (SEC, 2016, 2019b) has long discouraged boilerplate explanations, emphasizing the dissemination of relevant and material information to investors without using generic language. One reason for these institutional requirements is that lengthy risk factor disclosures could obstruct investors' understanding of the most significant risks (SEC, 2016, p.153). As risk management disclosure (RMD) is a part of risk factor disclosure, the SEC expressed concern that the inclusion of boilerplate risk management plans could further lengthen risk factor disclosures (SEC, 2016, p. 170). Many commenters also express disapproval of disclosures that contain generic risk factors, providing recommendations on how to reduce such boilerplate disclosures. For instance, Shearman & Sterling LLP (2014) suggest that Item 503(c) should include examples of generic disclosures that need not be included as risk factors. The United Kingdom's Financial Reporting Council (FRC, 2018) also has long sought to encourage entity-specific risk factor disclosure in corporate reporting due to concerns that boilerplate disclosures reduce the comprehensibility and decision-usefulness of annual reports.

Studies in psychology suggest that boilerplate disclosures can dilute important information, potentially weakening the usefulness of corporate reports for investors' decision-making (Zukier, 1982). Other research illustrates that boilerplate disclosures reduce investors' sensitivity to diagnostic information (Henry & Peytcheva, 2020) and influence their judgments (Arikan, 2021; Nelson & Pritchard, 2016).

However, there remains a gap between the regulatory requirements and actual practice. For example, while the CSRC requires risk management plans to be disclosed alongside risk factors in annual reports without using boilerplate language, many firms fail to provide such disclosures.[1] Moreover, many firms that disclose risk management plans use boilerplate language, whereas others provide more specific details (See Appendix A for RMD excerpts). Given companies' different choices in practice, it is important to determine the effects of RMD presence and specificity on investors' judgments. Tan and Yeo (2023) find that before a disclosed risk materializes, RMD only increases investors' valuation judgments when the corresponding risk is disclosed with a positive tone but has no effect when the risk is disclosed with a negative tone. We extend their research by finding that RMD specificity drives investors' judgments even when the risk is disclosed with a negative tone.[2]

[1] To illustrate this, we also randomly examined the 2023 annual reports of 20 A-listed food companies, and found that eight companies did not provide RMD.

[2] We have not investigated the impact of RMD during the post-materialization stage, as our focus is primarily on understanding, all other things being equal, how variations in the specificity of RMDs might influence investors' judgments. Intuitively, specific RMD could backfire once the disclosed risks materialize, as investors might perceive an increased severity of the risks and question the company's competence in managing them.

Scenario thinking theory suggests that individuals mentally simulate possible outcomes when making judgments in a given situation, and the ease with which they envision an outcome influences their expectations. Specifically, when an outcome is easier to imagine, people tend to perceive it as more likely to occur. Applying this to RMD, we argue that specificity influences investors' perceived effectiveness of the disclosed risk management plans. A more specific RMD makes it easier for investors to mentally simulate how the risks will be mitigated, reinforcing their belief that the measures will be effective. This, in turn, leads to more favorable investment judgments.

It is essential to investigate the effect of RMD specificity on investors' judgments for several reasons. First, regulatory bodies' attitudes toward RMD vary. The CSRC requires firms to disclose risk factors together with the risk management plans they have implemented or will implement. However, the SEC has expressed concerns regarding RMD. Specifically, Item 305b under the S-K Regulation requires firms to provide RMD, whereas Item 105 under the S-K Regulation does not allow risk management plans to be disclosed alongside risk factors. This is due to concerns that RMD would dilute investors' perceived seriousness of the corresponding risks (SEC, 2016). Second, the discrepancy between corporate practices and regulatory requirements highlights the need for companies to increase their emphasis on RMD, as our findings suggest. As recommended by standard setters, companies should prioritize the use of specific and detailed language over boilerplate terminology when providing RMD.

RMD specificity remains ill-defined, making it difficult to determine the appropriate indices without endogeneity concerns. To address this challenge, we use an experimental approach to investigate the relationship between RMD specificity and investor judgment. We use a 1×3 between-participants experimental design to test our hypotheses. Experiment 1 involves 96 MBA students who assume the roles of prospective investors. We manipulate RMD at three levels: (1) no disclosure of risk management plans, (2) disclosure of generic risk management plans and (3) disclosure of specific risk management plans. We find that, compared with no disclosure, specific RMD leads to more favorable investment judgments by non-professional investors. This relationship is sequentially mediated by investors' perceived specificity of the risk management plans, ease of mental simulation and perceived effectiveness of these plans. In contrast, disclosing generic risk management plans has little effect on investors' judgments. Experiment 2 provides further evidence and rules out alternative explanations that the observed effect is driven by either the length or format of RMD.

An examination of nonprofessional investors' judgments is particularly relevant because these investors represent a substantial portion of the capital market and significantly influence market dynamics. According to Shenwan Hongyuan Research Institute's classification of the A-share investor structure, retail investors were the largest investor category in China's A-share market in 2021, holding 34 % of all tradable shares,[3] and continued to account for 33.4 % of all investors in 2023.[4] Prior research also suggests that retail investors can influence overall market fluctuations (Barber et al., 2008; Aigbovo & Ilaboya, 2019). Given their significant impact, understanding how nonprofessional investors respond to RMD can provide valuable insights into market behavior and corporate disclosure strategies.

Our study makes both practical and theoretical contributions. Many disclosure standards require the disclosure of risk management plans alongside risk factors in annual reports and the use of specific, rather than boilerplate, language. However, not all companies provide RMD. For companies that do provide RMD, the disclosure language is often boilerplate. By demonstrating that the presence and specificity of RMD influence investors' judgments, we may help companies better understand the significance of disclosure in annual reports and inspire them to use specific (rather than boilerplate) language when providing RMD, as encouraged by standard setters. Second, while Tan and Yeo (2023) find that RMD does not influence investors' judgments when the tone of risk factors is negative, our research offers a different perspective. Tan and Yeo's results sug-

[3] The link to the related news is attached here for reference: https://www.thepaper.cn/newsDetail_forward_18606352.

[4] The link to the related news is attached here for reference: https://stock.finance.sina.com.cn/stock/go.php/vReport_Show/kind/lastest/rptid/778265203967/index.phtml.

gest that companies may benefit from disclosing RMD only when using a positive tone, but this poses a risk: as noted in prior research, optimistic tones in risk disclosures are linked to increased litigation risks (Rogers et al., 2011; Levy et al., 2018). Companies may face the dilemma of balancing the benefits of positive tones with the costs of potential litigation risks. In contrast, our findings show that even when a negative tone is used, specific RMD can still positively influence investors' perceptions. This finding suggests that companies can maintain transparency and provide decision-useful disclosures without relying on positive tones, potentially offering them a more sustainable and less risky strategy for managing risk disclosure. We believe that our findings have important implications not only for companies but also for regulators. Tan and Yeo (2023) also examine the effects of RMD effects in a U.S.-centric context, where specificity in RMD is not highlighted, and thus potentially overlook its significance in jurisdictions such as China. The CSRC's explicit focus on specificity reflects a broader push to prioritize decision-useful information, which our study directly addresses. By testing how the specificity of RMD influences investors' judgments, we offer a perspective that Tan and Yeo's framework does not fully capture.

Third, our study supplements the risk disclosure literature by exploring the effect of RMD specificity. Prior studies mostly focus on the specificity of risk factor disclosures (e.g., Kravet & Muslu, 2013; Campbell et al., 2014; Cazier et al., 2021; Arikan, 2021; Zhao et al., 2022) but do not consider the specificity of RMD. Certainly, companies providing detailed risk factors may also offer more specific RMD, which would enhance the decision-usefulness of their disclosures. Therefore, we aim to fill this gap by examining the effect of RMD specificity on investment judgments, thereby contributing new insights to the literature on the decision-usefulness of specific disclosures. We also provide a mechanism based on scenario thinking theory to explain how this effect occurs: investors perceive that risk management plans including specific RMD are more effective, which leads to more favorable investment judgments.

The remainder of this paper is organized as follows. Section 2 describes the theory and hypotheses. Sections 3 and 4 outline our design and the results of Experiment 1, respectively. Sections 5 and 6 outline the research design and results of Experiment 2, respectively. Finally, Section 7 presents our conclusions.

2. Theory and hypothesis Development

2.1. The effect of RMD in annual reports on Investors' judgments

Firms disclose both financial and non-financial information in their annual reports to provide a comprehensive view of their performance. Research suggests that in addition to financial disclosures, non-financial disclosures provide indispensable supplementary information that helps build investors' confidence and encourages investment in the disclosing firm (Beattie et al., 2004; Li et al., 2024; Liu et al., 2025). As a type of non-financial disclosure, RMD introduces the measures taken to mitigate disclosed risks and proves the effects on essential constructs such as firm value (Abdullah et al., 2015). One question arises: what happens when companies do not provide RMD? Tan and Yeo (2023) find that RMD influences investors' judgments only when the corresponding risk is disclosed with a positive tone; in contrast, it has no effect when the tone is negative. However, their study used generic language when providing RMD. It is unclear whether providing detailed and specific RMD could help firms establish credibility among investors, regardless of the risk disclosure tone.

2.2. Scenario thinking theory and the effect of RMD specificity on investors' judgments

Scenario thinking theory includes a family of psychological theories that discuss the influences of information structures and imagination on individuals' judgments (cf. Sedor, 2002). Scenario thinking theory mainly involves heuristic availability (Tversky & Kahneman, 1973; Kahneman & Tversky, 1982), explanation (Koehler, 1991) and mental simulation during scenario thinking (Schoemaker, 1991). Heuristic availability theory suggests that people tend to make judgments based on the information they already have about a situation (i.e., information that is highly available). Regarding explanation, Koehler (1991) observes that after a person explains or imagines the possibility of an event occurring, their confidence that the probable event will become a reality is increased because they had temporarily assumed the hypothesis regarding the event's occurrence to be true when they explained or imagined it. Schoemaker (1991) defines a scenario as a script-like characterization of a possible future, which is presented in considerable detail with particular emphasis on causal connections, internal consistency and concreteness. Overall, scenario thinking theory suggests that when judging an uncertain situation, individuals mentally simulate a sequence of events from the initial state to the outcome. Their judgments are influenced by the ease of this mental simulation: the easier it is to envision the process, the more likely they will perceive the event to be, and vice versa.

There are precedents for applying scenario thinking theory to accounting research. Sedor (2002) examines whether different presentations of managers' future plans influence analysts' forecast optimism. The author finds that analysts are more optimistic when managers use a scenario-inducing format than when they use a simple list. Similarly, Leitter et al. (2022) examine the effects of separately identifying intangible assets and providing strategic disclosures on investors' judgments. They find that either separately identifying intangible assets or providing a strategic disclosure makes it easier for investors to mentally simulate successful M&A outcomes, leading to more favorable investment judgments. We also use scenario thinking theory to develop our hypothesis. We expect that when investors are confronted with risk disclosure alongside risk management plans, they will mentally simulate how well the plans can mitigate the risks. During this process, the initial state is when the risk is disclosed, and the final state is when the risk is realized (or, conversely, when the risk management plans mitigate the corresponding risk as expected). According to scenario thinking theory, investors' judgments depend on how easily they can simulate this process.

There are two forms of boilerplate disclosures: vertical boilerplate, where a company repeats its prior-period disclosures, and horizontal boilerplate, where multiple companies use essentially identical language (Henry & Peytcheva, 2020). In this paper, we focus on the effect of horizontal boilerplate language in RMD.

The specificity of RMD is ill-defined. However, as prior studies address boilerplate language in risk factor disclosures either archivally or experimentally, we can use this as a reference when measuring the specificity of RMD. Archival studies report two main types of measures of boilerplate risk disclosure. The first measure involves capturing the degree of language standardization (Lang & Stice-Lawrence, 2015); that is, four-word phrases commonly used in risk disclosures are identified, and the frequency of occurrence of these four-word phrases in the sample companies' annual reports is measured. However, this approach is unsuitable for our study as it does not directly capture the specificity of disclosures. Instead, we adopt the second method, which defines specificity more directly by identifying disclosures containing concrete details, such as the names of persons, organizations and locations; quantitative values in percentages; monetary values in dollars; times and dates (Hope et al., 2016). Few experimental studies explore risk disclosure specificity. Arikan (2021) investigates the direct impact of risk factor disclosure specificity on investors' judgments. He defines specificity in two dimensions: vividness and signaling. Vividness refers to the fact that, compared with generic disclosures, specific risk disclosures are often more vivid because they cite real-life examples and use more graphic language. Signaling refers to the fact that companies tend to use more specific language in their announcements when they face more serious risks, and investors notice such signals (Kasznik & Lev, 1995).

Combining scenario thinking theory with the methods used in previous studies on boilerplate risk factor disclosure, we define specific RMD as a disclosure that explains concrete risk management strategies in detail by incorporating elements such as named individuals, organizations, locations, quantitative values (e.g., percentages and monetary amounts), numerical frequencies (e.g., occurrence counts) and time-related details (e.g., dates and durations). By providing such detailed information, specific RMD enhances the reader's ability to envision the implementation of these strategies and assess their potential effectiveness in mitigating the corresponding risks. We expect that when investors are confronted with risk factors and RMDs, they mentally simulate whether the plans will be implemented and can mitigate the corresponding risks. Both generic and specific RMD provide investors with resources for mental simulation. However, specific RMD makes this process easier by leading investors to perceive a higher likelihood that the risk management plans will work as expected. In other words, specific RMD may enhance investors' perception of the plans' effectiveness. This perception can lead investors to decrease their judgments of risks and thus make more favorable investment judgments (Huber et al., 2019). Therefore, when faced with the same risk, specific RMD can increase investors' willingness to invest by improving their judgment of its effectiveness to a greater extent than generic RMD. Accordingly, we propose H1.

H1: Specific RMD improves investors' perceived effectiveness of the disclosed risk management plans, leading to more favorable investment judgments, while generic RMD has a weaker influence on investors' judgments.

Our hypothesis does not lack tension. Arikan (2021) finds that specific risk factors signal the seriousness of the disclosed risks, which may decrease the likelihood that nonprofessional investors will invest in a company. Therefore, when provided with specific RMD, investors may perceive the corresponding risks as more substantial. They may also expect that the disclosed risk management plans will be less likely to mitigate the risks, leading to less favorable investment judgments. According to Arikan (2021), however, it is very hard to control the nature of the risk itself. In our experimental setting, we hold the disclosed risk factors constant and only vary the specificity of the risk management plans.

3. Experiment 1

3.1. Participants

Our sample comprises 96 MBA students from a Chinese university who received course credit for participating in the study, which required approximately 15 min to complete. According to prior studies, MBA students are appropriate proxies for nonprofessional investors (Libby et al., 2002; Hirst et al., 2007; He et al., 2019). Therefore, we chose not to use participants more sophisticated than necessary to achieve our experimental goals. The average age of the participants was 30.04 years. Forty-three participants (44.8 %) were female. The participants had an average of 7.4 years of full-time work experience and had taken an average of 3.56 accounting and finance courses. Fifty-nine participants (61.5 %) had stock investment experience, with a mean duration of 4.14 years. Eighty-nine participants (92.7 %) had experience in reading annual reports.

3.2. Design

We use a 1×3 between-participants experimental design to investigate the effect of RMD on investors' judgments. The manipulation involves three conditions: non-disclosure, generic disclosure and specific disclosure of risk management plans. The dependent variable is investment judgment, measured by averaging the participants' ratings of the attractiveness of the investment and their willingness to invest. The potential mediators are the investors' perception of the specificity of the risk management plans, ease of mental simulation and perception of the effectiveness of the risk management plans. We hypothesize that compared with the *non-disclosure* condition, the presence of RMD (specific or generic) leads investors to perceive the effectiveness of the risk management plans more positively; in turn, they will make more favorable investment judgments. In addition, we expect to find that participants in the *specific* RMD condition perceive higher specificity and find it easier to mentally simulate the implementation of the disclosed plans. This, in turn, should increase the perceived effectiveness of the risk management plans and ultimately lead to more favorable investment judgments than those made by participants in the *generic* group. Our expected model is demonstrated in Fig. 1.

3.3. Materials, manipulations and procedure

Our experimental materials are based on a food company listed on the Hong Kong Exchanges and Clearing Limited (HKEX). We first provided the company's background information and historical financial data. Subsequently, we provided the company's risk disclosure information. In the *non-disclosure* condition, participants were provided with only risk factor disclosures. In the other two conditions, RMDs were presented just

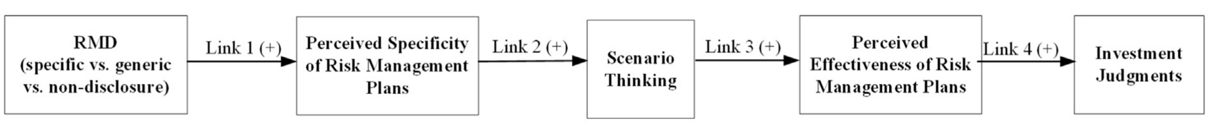


Fig. 1. Conceptual model of the effects of RMD on investment judgments, comparing specific RMD vs. generic RMD vs. non-disclosure. The parenthetical comment next to each link represents the expected coefficient sign.

below the corresponding risks. In the *specific* and *generic* conditions, we provided specific and generic risk management plans, respectively. Appendix B details our specificity manipulation.

We used a negative tone when describing the risk factors to echo the findings of Tan and Yeo (2023). More importantly, archival evidence shows that in practice, companies use negative tone more often than positive tone when disclosing risk factors (Tan & Yeo, 2023).[5] For example, the food safety risk is expressed as "once the Company experiences food safety problems...it may face product liability claims, a negative brand impact or government penalties, which may adversely affect the Company's reputation and cause a loss of economic interests," rather than "the Company will ensure food safety...thus protecting itself from product liability claims and government penalties and preserving the brand image, which in turn will preserve the Company's reputation and enhance its economic interests." As Tan and Yeo (2023), who provide generic RMD in their study, find that RMD alongside the disclosure of risk factors with a negative tone has no effect on investors' judgments, we aim to extend their findings by exploring the use of specific (vs. generic boilerplate) language when disclosing risk management plans.

The second key design feature is specific RMD vs. generic RMD. Holding the subtitle of the risk management plans constant, we provided further detailed explanations in the *specific* condition. For example, regarding food safety risk (see Appendix B), the first measure in the *generic* condition was "optimizing its food safety testing methods"; in contrast, in the *specific* condition, details were provided to explain further how the company optimizes its food safety testing methods. These details include specific information such as what hardware facility the company prepares (e.g., a high-standard testing center equipped with top international testing facilities) to optimize food safety testing methods and what testing results the company has obtained in the past year (e.g., over 1,400 batches of sampling inspection were implemented, and the pass rate was 100 %). The differences in information content under the two conditions may raise concerns. First, boilerplate language, by nature, tends to be less informative than specific language (Arikan, 2021). Second, the absence of specific risk management plan disclosure does not necessarily imply that the company lacks the capability to implement such plans or is not already doing so. In our comparison of several companies in the food industry with similar market values and net profits, we found variations in the disclosure practices: some companies opted for boilerplate RMDs, while others provided more specific RMDs. Therefore, our design not only enhances the prominence of RMD specificity but also aligns closely with real-world practices. Furthermore, we ensured that the readability of the RMD was the same between the *generic* and *specific* conditions. Specifically, we asked the participants to self-report the extent to which they found the RMD to be easy to understand. We find no significant difference in this metric between the two conditions (7.39 in the *generic* condition vs. 7.50 in the *specific* condition, $p = 0.854$, untabulated).

Another key design feature is the quantification of specific RMD manipulation. To construct our manipulation, we referenced the framework established by Hope et al. (2016), who define specificity in risk factor disclosure as encompassing "the number of specific entity names, quantitative values in percentages, monetary values in dollars, timeframes, and dates." Based on their validated approach, we believe that including quantification is not only consistent with the construct of specificity but also adds clarity to our manipulation in a meaningful way.[6] We acknowledge concerns that quantification alone could drive our results but argue that quantitative information in RMD helps investors better understand a company's future risks and strategies and affects investors' judgment only when it is relevant to the accompanying qualitative details. This notion is supported by the CSRC, which highlights that quantitative data should enhance transparency in a way that

[5] We also review the risk factor disclosures of the top 10 companies in the food industry. Upon visual inspection, a negative tone is used much more frequently than a positive tone.

[6] Additionally, the inclusion of quantification in specific RMD follows established reporting practices used by some companies in their risk management disclosures. For example, Dali Foods Group Company, in its 2020 annual report, provides detailed quantitative information about its food safety risk management plan, stating that "In 2020, a total of more than 2,000 internal and external random inspections were conducted, with a passing rate of 100%." Similarly, Uni-President China's 2023 annual report states that "As of 2023, 173 projects of the Group have been approved. Meanwhile, the Group has a number of national patents for utility models and more than 700 perennial independent testing projects" when discussing its risk management plans against food safety risk. These companies' actual disclosures demonstrate that our approach to including quantification in specific RMD reflects the current practices of corporate disclosure and enhances the external validity of our manipulation.

Table 1
Descriptive statistics and test of Hypothesis 1 (Experiment 1).

Panel A: Descriptive Statistics (dependent variable: investment judgments)

RMD	N	Mean	Std. Deviation	Std. Error	Min	Max
Specific	33	6.27	1.640	0.285	2.50	9.00
Generic	32	5.31	2.090	0.369	1.00	9.00
Non-disclosure	31	5.06	2.093	0.376	0.50	9.00
Total	96	5.56	1.998	0.204	0.50	9.00

Panel B: One-way ANOVA

Source	Sum of Squares	df	Mean Square	F	p-value
Between Groups	26.334	2	13.167	3.471	0.035
Within Groups	352.791	93	3.793		

Panel C: Post Hoc Tests (LSD)

				95 % Confidence Interval	
Source	Mean Difference	Std. Error	p-value	Lower Bound	Upper Bound
Specific vs. Generic	0.960	0.483	0.050	0.001	1.92
Specific vs. Non-disclosure	1.208	0.487	0.015	0.24	2.18
Generic vs. Non-disclosure	0.248	0.491	0.615	−0.73	1.22

informs investors about the company’s outlook (CSRC, 2021). Not all quantitative information represents specificity; instead, specificity refers to the meaningful integration of these elements into a cohesive narrative.

The participants were randomly assigned to one of three conditions and told to assume the role of a general investor considering investing in A&M, a hypothetical Chinese company. First, they were asked to read the background information explaining that A&M is a listed company engaged in the R&D, production and retail sale of food and drinks. Then, they read the financial data selected from A&M’s past three annual reports (fiscal years 2019–2021), showing growth in 2020 and a decline in 2021. Specifically, the total assets and operating revenue increased from 2019 to 2021, as did the operating cost. The gross profit, net profit, net cash flows from operating activities and EPS were slightly higher in 2021 than in 2019 but remained lower than those in 2020.[7]

After reading these materials, the participants responded to a series of questions, including (1) measurement of the dependent variables, (2) measurement of the mediator variables, (3) manipulation checks and (4) demographic questions. To measure participants’ investment judgments, we asked them to evaluate (1) the attractiveness of the investment on an 11-point scale on which 0 = “not at all attractive” and 10 = “extremely attractive” and (2) their willingness to invest on an 11-point scale on which 0 = “absolutely not willing to invest” and 10 = “absolutely willing to invest.” Inspired by Leitter et al. (2022), we then measured the ease of the participants’ mental simulation by asking them to what extent they agreed with two statements on an 11-point scale: (1) “I think it is easy to imagine the process of implementing the risk management plans at A&M,” and (2) “I think it is easy to imagine that A&M successfully mitigates the disclosed risks as an outcome.” Next, the participants responded to questions about their perception of the effectiveness of the risk management plans (if provided with RMD). Specifically, we captured the participants’ perceptions regarding three aspects: perceived feasibility of the risk management plans, perceived probability of implementation of the risk management plans and perceived probability that the risk management plans would successfully mitigate the disclosed risks if implemented. Finally, the participants answered the manipulation checks and demographic questions.

[7] We do not design the financial performance as steadily increasing for two reasons. First, prior research suggests that good performance may lead to a ceilings effect (Koonce et al., 2015), thus leaving little room for incremental positive effects of specific RMD to be detected. Second, considering the prominently negative impact COVID has had on the food industry over the past few years, it is almost unrealistic for a food company to have an increasing trend in its financial performance.

4. Results of experiment 1

4.1. Manipulation checks

To test the manipulation of RMD specificity, we asked the participants whether the risk management plans were disclosed in the case material provided. If their answer was "yes," we asked them to evaluate to what extent the disclosed risk management plans were specific using an 11-point scale, with 0 = "very low" and 10 = "very high." If their answer was "no," then they did not need to answer this question. Seventy-five of the 96 participants (78.13 %) answered the question correctly. Specifically, the correct answer was given by 28 of 33 participants (84.85 %) in the *specific* condition, 23 of 32 participants (71.88 %) in the *generic* condition and 24 of 31 participants (77.42 %) in the *non-disclosure* condition, with one participant failing to respond to the manipulation check questions. The correct response rate was lowest in the *generic* condition, indicating that boilerplate disclosures are more likely to be neglected than specific disclosures. Among the participants who correctly answered the first question, those in the *specific* condition rated the specificity as significantly higher than those in the *generic* condition (5.59 vs. 3.96, $F_{1,49} = 5.627$, two-tailed $p = 0.022$, untabulated), indicating successful manipulation. Our statistical results do not vary significantly when excluding participants who failed the manipulation check question. Therefore, we retain these participants in our subsequent analyses.

4.2. Hypothesis testing

H1 predicts that relative to non-disclosure, disclosing specific risk management plans will lead to more favorable investment judgments due to higher perceived effectiveness of the disclosed plans, whereas providing generic risk management plans will not influence investors' investment decisions.

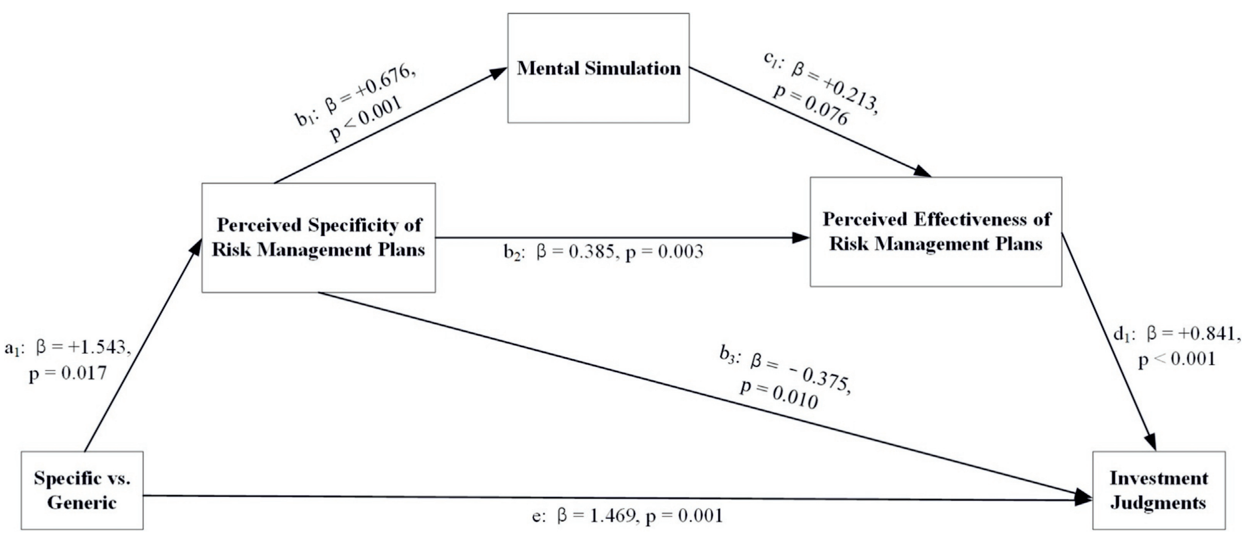


Fig. 2. **Path analysis for Experiment 1** based on two-tailed tests[1]. The standardized coefficients and corresponding p-values are shown next to each link[2]. Mental simulation is a latent variable inferred from the measurements of two observable variables: mental simulation of implementation and mental simulation of successful mitigation. Perceived effectiveness of risk management plans is a latent variable inferred from the measurements of three observable variables: perceived feasibility of risk management plans, perceived probability of implementation of risk management plans and perceived probability that risk management plans will successfully mitigate the disclosed risks if implemented. The overall model fit indices are as follows: (a) $\chi^2 = 15.490$, df = 13, CMIN/DF = 1.192 and p = 0.278; (b) RMSEA = 0.062, below the cut-off point of 0.08 for a good fit and (c) SRMR = 0.058, below the cut-off point of 0.08 for a good fit. (d) The incremental fit indices are GFI = 93.5 %, TLI = 96.5 %, NFI = 91.5 % and CFI = 98.4 %. [1] We excluded participants who did not respond to the specificity of the risk management plans, as this measure is necessary for the first mediator in our model. [2] Paths that are not significant are not depicted in the model, and the corresponding p-values for these insignificant paths are omitted as well.

Table 1, panel A reports the descriptive statistics of the participants' investment judgments under the three conditions. Panel B presents the one-way analysis of variance (ANOVA) results, which suggest a significant effect of RMD specificity on participants' investment judgments ($F_{2,93} = 3.471$, two-tailed $p = 0.035$). The post hoc test result in Panel C shows a significantly higher investment judgment score for the *specific* group than for the *generic* (mean difference = 0.960, two-tailed $p = 0.050$) and *non-disclosure* groups (mean difference = 1.208, two-tailed $p = 0.015$). However, there is no significant difference between the means of the *non-disclosure* and *generic* groups, with Tan and Yeo (2023)'s findings. These results imply that when risk factors are disclosed using a negative tone, generic RMD does not influence investors' judgments. Only specific RMD can result in more favorable investment judgments.

To test our proposed mechanism, we conduct a structural equation analysis with the perceived specificity of risk management plans, mental simulation and perceived effectiveness of risk management plans as the sequential mediators. As shown in Fig. 2, higher specificity leads to a higher perceived specificity of risk management plans (coefficient = 1.543, $p = 0.017$), which in turn leads to higher mental simulation (coefficient = 0.676, $p < 0.001$) and a higher perceived effectiveness of risk management plans (coefficient = 0.213, $p = 0.076$). Finally, these lead to more favorable investment judgments (coefficient = 0.841, $p < 0.001$).

4.3. Alternative explanations

While the mediation analyses show that specific RMD leads to more favorable investment judgments driven by scenario thinking than does generic RMD, there remain some alternative explanations. First, specific RMD could provide cues about management competence. To rule out this alternative explanation, we asked the participants to indicate to what extent they agreed with the following statements, using an 11-point scale (–5 = "Not at all agree," 5 = "Very strongly agree"): (1) "I believe that A&M's management is very competent" and (2) "I think A&M's management is very trustworthy." One-way ANOVA shows that the participants' ratings of management competence and trustworthiness were not higher in the *specific* condition than in either the *generic* condition (competence: 5.67 versus 5.78, two-tailed $p = 0.802$; trustworthiness: 5.76 versus 5.63, two-tailed $p = 0.776$, untabulated) or the *non-disclosure* condition (competence: 5.67 versus 5.35, two-tailed $p = 0.499$; trustworthiness: 5.76 versus 5.45, two-tailed $p = 0.515$, untabulated). Second, RMD specificity might influence investors' perception of the reliability of the firm's disclosure, which in turn would affect their investment judgments. Therefore, we asked the participants to indicate to the extent to which they agreed with the following statement, using an 11-point scale (–5 = "Not at all agree," 5 = "Very strongly agree"): "I believe that A&M's disclosure reliability is very high." One-way ANOVA shows that the participants' scores for this question were not higher in the *specific* condition than in either the *generic* condition (5.85 versus 5.78, two-tailed $p = 0.887$, untabulated) or the *non-disclosure* condition (5.85 versus 5.74, two-tailed $p = 0.824$, untabulated). A third alternative explanation is that more specific RMD may lead to more favorable investment judgments because investors' perception of the risk severity may be diluted. We perform an additional one-way ANOVA, using participants' risk perception as the dependent variable. We measure risk perception by averaging the participants' responses to the following two questions: (1) "How high do you think A&M's risk level is?" (0 = Very low to 10 = Very high), and (2) "How risky do you consider an investment in A&M's stock to be?" (0 = Not at all risky to 10 = Extremely risky). The untabulated results show that participants in the *specific* RMD condition did not perceive a lower risk level than those in the *generic* condition (5.39 versus 5.53, two-tailed $p = 0.735$) or the *non-disclosure* condition (5.39 versus 5.89, two-tailed $p = 0.229$). Overall, these further analyses provide no evidence supporting the alternative explanations. Consequently, RMD specificity leads to more favorable investment judgments, and this relationship is driven by investors' scenario thinking.

5. Experiment 2

5.1. Motivation

We conducted a follow-up experiment for two key reasons. First, in Experiment 1, we manipulated RMD specificity by adding detailed explanations to the *specific* condition. However, the number of words in the *specific* condition far outweighed that in the *generic* condition, raising concerns that our result was driven by the length of the disclosures, rather than the content. To address this concern, we ensured that the word counts of the disclosures used in Experiment 2 are as similar as possible. Second, we used ordered lists to present detailed explanations in the *specific* condition in Experiment 1. In its *Plain English Handbook*, the SEC (1998) suggests that the use of formatting features such as tables, bullet points and listed numbers increases the legibility and readability of a disclosure. Accounting research shows that how information is presented can affect investors' perceptions and decision-making (Rennekamp, 2012). Therefore, we designed two forms of generic RMD in Experiment 2, with and without listed numbers, to test whether the presentation and format of disclosures would affect investors' judgments.

5.2. Participants

Our participant sample comprises 44 MBA students from the same university as in Experiment 1; these students also received course credits for participating in the study. We excluded one participant who failed to complete the study. For the remaining 43 participants, the average time to complete the study was approximately 15 min. The average age of the participants was 31.86 years. Twenty-eight participants (65.1 %) were female. The participants had an average of 9.1 years of full-time work experience and had taken an average of 2.83 accounting and finance courses. Thirty-six participants (83.7 %) had stock investment experience, with a mean duration of 5.31 years. Forty-one participants (95.3 %) had experience in reading annual reports.

5.3. Design

We use a $1 \times 2 + 1$ between-participants design and manipulate RMD specificity (*specific* vs. *generic-listed*). We also include a *generic-blocklike* condition to examine whether the format of ordered lists affects investors' judgment. We exclude the *non-disclosure* condition because the results from Experiment 1 show no difference between non-disclosure and generic RMD. The dependent variable and potential mediators are the same as in Experiment 1.

5.4. Materials, manipulations and procedure

We set the same experimental instructions, background information and historical financial data as in Experiment 1 and modify the risk management plans. Specifically, we refer to the definition of specificity in

Table 2
Descriptive Statistics and Test of Hypothesis 1 (Experiment 2).

Panel A: Descriptive Statistics (dependent variable: investment judgments)

RMD	N	Mean	Std. Deviation	Std. Error	Min	Max
Specific	14	7.29	1.602	0.428	4.50	10.00
Generic-listed	15	5.10	2.189	0.565	0.50	7.50
Generic-blocklike	14	5.00	1.912	0.511	1.00	7.50
Total	43	5.78	2.156	0.329	0.50	10.00

Panel B: One-way ANOVA

Source	Sum of Squares	df	Mean Square	F	p-value
Between Groups	47.194	2	23.597	6.379	0.004
Within Groups	147.957	40	3.699		

Panel C: Post Hoc Tests (LSD)

				95 % Confidence Interval	
Source	Mean Difference	Std. Error	p-value	Lower Bound	Upper Bound
Specific vs. Generic-listed	2.186	0.715	0.004	0.74	3.63
Specific vs. Generic-blocklike	2.286	0.727	0.003	0.82	3.75
Generic-listed vs. Generic-blocklike	0.100	0.715	0.889	−1.34	1.54

Hope et al. (2016) and Arikan (2021) and add examples, quantitative values in percentages, money values in dollars, times and dates to the *specific* condition. For example, in the *specific* condition, the disclosure includes "introducing advanced testing facilities and technology. In 2022, we invested up to USD 10 million for the continuous evaluation and improvement of our testing process on a quarterly basis"; the counterpart in the *generic-listed* condition is "introducing advanced testing facilities and technology and evaluating and improving our testing processes on a regular basis." The *generic-blocklike* condition is the same as the *generic-listed* condition except that does not include listed numbers. See Appendix C for more details on our manipulation of specificity.

6. Results of experiment 2

6.1. Manipulation checks

The manipulation check questions are the same as in Experiment 1. Thirty-four of 43 participants (79.07 %) answered the true–false question correctly. Specifically, 13 of 14 participants (92.86 %) in the *specific* condition, 13 of 15 participants (86.67 %) in the *generic-listed* condition and seven of 14 participants (50 %) in the *generic-blocklike* condition answered the question correctly. The low pass rate in the *generic-blocklike* condition suggests that blocklike boilerplate disclosures are more likely to be neglected than the other disclosure types. We speculate that participants in this condition noticed the RMD but did not consider it to include risk management plans due to its generic nature. Among the participants with correct answers to the first question, the specificity was rated as significantly higher among those in the *specific* condition than among those in the *generic-listed* (6.08 versus 4.46, two-tailed $p = 0.067$, untabulated) and *generic-blocklike* conditions (6.08 versus 3.25, two-tailed $p = 0.007$, untabulated); however, there was no significant difference between the *generic-listed* and *generic-blocklike* conditions (4.46 versus 3.25, two-tailed $p = 0.223$), indicating successful manipulation. Our statistical inferences remain unchanged when we exclude participants who failed to correctly answer the question. Therefore, we include all participants in our subsequent analyses.

6.2. Hypothesis testing

Table 2, panel A reports the descriptive statistics of the participants' investment judgments under the three conditions. Panel B presents the results of one-way ANOVA, which suggest that RMD specificity significantly affects participants' investment judgments ($F_{2,40} = 6.379$, two-tailed $p = 0.004$). The post hoc test result in Panel C shows that investment judgments are significantly higher in the *specific* group than in either the *generic-listed* (mean difference = 2.186, two-tailed $p = 0.004$) or *generic-blocklike* group (mean difference = 2.286, two-tailed $p = 0.003$). There is no significant difference between the means of the *generic-listed* and *generic-blocklike* groups, indicating that the use of listed numbers does not change investors' perception of the effectiveness of risk management plans.[8]

To verify our proposed mechanism, we conduct a structural equation analysis, with perceived specificity of risk management plans, mental simulation and perceived effectiveness of risk management plans as the sequential mediators. As shown in Fig. 3, higher specificity leads to a higher perceived specificity of risk management plans (coefficient = 2.077, $p = 0.006$), which in turn leads to higher mental simulation (coefficient = 0.285,

[8] There are two possible explanations for this result. First, While prior research suggests that the presentation form of disclosures is one aspect of readability (Rennekamp, 2012), it is possible that merely altering the presentation form does not meaningfully affect investors' perceived readability of the disclosures. Prior research on readability often manipulates multiple dimensions beyond presentation format, such as whether using difficult language, long sentences and passive voice (Rennekamp, 2012; Tan, Wang, & Zhou, 2014; Asay, Elliott, & Rennekamp, 2017). Given this, our findings suggest that presentation form alone may not be a strong enough factor to influence how investors assess the specificity of RMD. Second, even if the presentation form does impact readability, our results suggest that investors do not heavily rely on it when forming judgments about the effectiveness of risk management plans. It is likely that in the context of risk management disclosures, investors may prioritize substantive content over stylistic features such as presentation format. This implies that while readability can play a role in shaping perceptions of narrative disclosures, its impact may be less pronounced when investors focus on assessing the effectiveness of risk management strategies, especially when the involved risks of the food company is more familiar to investors.

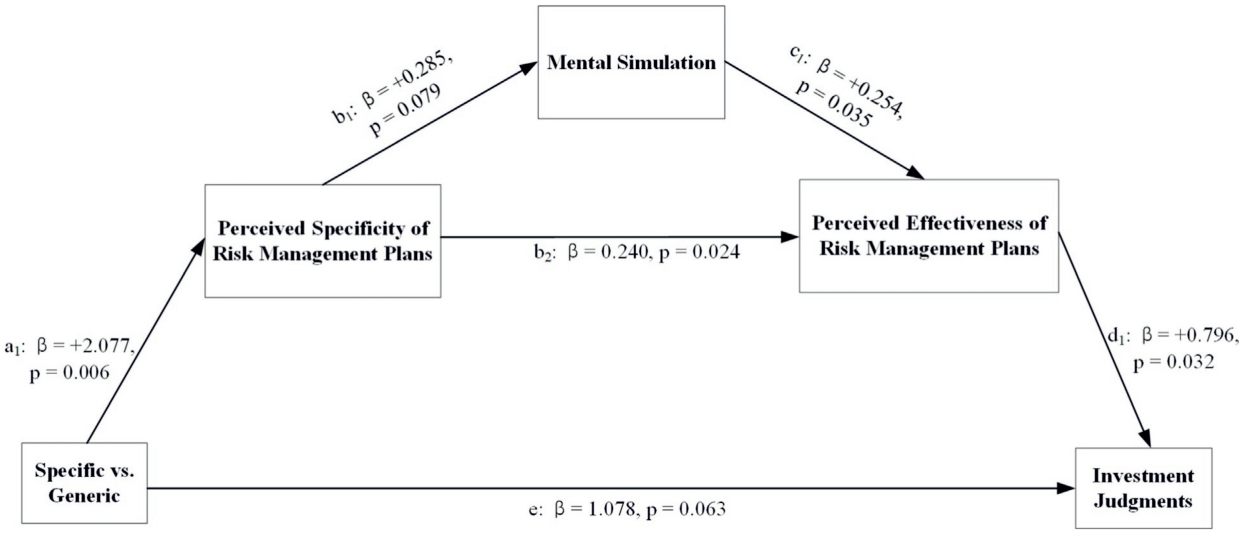


Fig. 3. **Path analysis for Experiment 2** based on two-tailed tests[1]. The standardized coefficients and corresponding p-values are shown next to each link[2]. Mental simulation is a latent variable inferred from measurements of two observable variables (as in Experiment 1): mental simulation of implementation and mental simulation of successful mitigation. Perceived effectiveness of risk management plans is a latent variable inferred from measurements of three observable variables (as in Experiment 1): perceived feasibility of risk management plans, perceived probability of implementation of risk management plans and perceived probability that risk management plans will successfully mitigate the disclosed risks if implemented. The overall model fit indices are as follows: (a) $\chi^2 = 5.435$, df = 13, CMIN/DF = 0.418 and p = 0.964; (b) RMSEA = 0.000, below the cut-off point of 0.08 for a good fit and (c) SRMR = 0.038, below the cut-off point of 0.08 for a good fit. (d) The incremental fit indices are GFI = 96.3 %, TLI = 100 %, NFI = 95.4 % and CFI = 100 %. [1] As in Experiment 1, participants who did not respond to the specificity of the risk management plans were excluded because this measure is necessary for the first mediator in our model. [2] As in Experiment 1, paths that are not significant are not depicted in the model, and the corresponding p-values for these insignificant paths are omitted as well.

p = 0.079) and a higher perceived effectiveness of risk management plans (coefficient = 0.254, p = 0.035). Finally, these lead to more favorable investment judgments (coefficient = 0.796, p = 0.032).

7. Discussion

In this study, we investigate the impact of the presence and specificity of RMD on investors' judgments through two experiments. We find that relative to generic RMD, specific RMD leads to a higher perception of the effectiveness of risk management plans, which increases the favorability of non-professional investors' judgments. We also find no difference between the generic RMD and non-disclosure conditions in terms of investors' judgments. The results in Experiment 2 further show that varying the presentation of generic RMD does not alter investors' perceived effectiveness of the risk management plans or their investment judgments.

In practice, many companies do not disclose risk management plans in their annual reports, despite the requirement to do so. Companies that disclose such plans often use boilerplate language, which fails to provide decision-useful information to investors. Based on scenario thinking theory, we argue that specific RMD can simplify the process of mental simulation for investors, thereby enhancing their perceived effectiveness of the plans in terms of mitigating the corresponding risks. Consequently, the investors' willingness to invest in the company is increased. Our findings should encourage companies to disclose specific risk management plans in accordance with regulatory standards, given the minimal impact of generic RMD on investors' decisions.

Tan and Yeo (2023) find that RMD influences investors' judgments only when the corresponding risk is disclosed using a positive tone. However, other research suggests that optimistic tones in risk disclosures can increase litigation risks (Rogers et al., 2011; Levy et al., 2018). This creates a dilemma for firms as they must choose between using a positive tone to maximize the benefits of RMD while facing potential legal consequences or using the more common negative tone and losing the advantages of RMD. Our findings provide a solution to this challenge by demonstrating that specific RMD enhances investors' perceptions even when the risk factors are framed negatively. In other words, firms can improve their disclosure transparency and provide decision-useful information without the added risks associated with positive tones. This insight is particularly important as many companies currently disclose generic risk management plans alongside negatively framed risk factors—a practice that seems to be fruitless effort.

Our findings also extend the literature on boilerplate disclosures. Prior studies focus on the influences of risk factor disclosure using boilerplate language (Arikan, 2021; Henry & Peytcheva, 2020; Cazier et al., 2021). Instead, we examine RMD, which is more aligned with practice as in many regimes, RMDs are required and often provided alongside risk factor disclosures in companies' annual reports.

Our study has some limitations that can be addressed in future research. First, we focus on the effect of RMD on non-professionals' investment judgments, but this effect may differ for professional investors who have more financial expertise. Future research could examine how professionals interpret and respond to RMD to provide a more comprehensive understanding of its effects. Second, we use a hypothetical food company in our experiments; while allowing for experimental control, this may not fully capture the complexities of investors' decision-making in the real world. The effects may also vary across industries, particularly when investors are less familiar with the associated risks. Future research could explore how industry-specific factors influence the relationship between RMD specificity and investors' judgments. Finally, we examine RMD within the context of annual reports. However, managers may communicate risk management plans through other channels, such as conference calls or press releases, which differ in format and investor engagement. Future research could explore whether the medium of disclosure affects how investors respond to RMD.

Data availability

Contact the authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Wei Chen, Jun Huang, Peishu Sheng, Yifei Xia, Shuang Xue, Yao Yu, Ziye Zhao, and seminar participants at Shanghai University of Finance and Economics for their helpful comments. The work described in this paper was supported by the National Natural Science Foundation of China (Grant Nos. 71972120 and 72172083), National Social Science Fund of China (Grant No. 21BGL083), and the Major Project of the Key Research Base for Humanities and Social Science of Ministry of Education of China (Grant Nos. 15JJD630008 and 22JJD790094).

Appendix A. Examples of risk-related disclosure in annual reports[9]

1. [non-disclosure condition] Excerpt from S&P International's annual report (2023)

Price fluctuation in and shortage of raw materials and perishability of coconuts may materially and adversely affect business operations

The coconut-related food manufacturing industry depends on a sufficient supply of major raw materials, namely coconuts and white kernels, at commercially reasonably prices. If our suppliers for any particular raw material are unable or unwilling to meet our requirements, we could suffer shortages or significant cost

[9] In brackets are the corresponding specificity conditions. The risk management plans disclosed by companies are underlined in the examples provided.

increases. In addition, any shortage or disruption in our supply of coconuts and other raw materials could affect our performance and our ability to satisfy the purchase orders of our customers, which may adversely affect our profitability, results of operations and financial condition.

Coconut is a perishable raw material which may deteriorate due to delivery delays, or poor handling during transportation by suppliers or logistic partners. This may result in failures to operate production of coconut products, thereby damaging our business and/or reputation. If any raw materials or finished products are alleged or found to be spoiled, contaminated, tampered with, incorrectly labelled, unsafe or otherwise associated with food safety incidents, we could be subject to product liability claims, adverse publicity and regulatory investigation, intervention or penalties or product returns, any of which may result in decreased profitability as well as damage to our brands and reputation.

2. [*generic* condition] *Excerpt from DaChan Food (Asia)'s annual report (2023)*

Food safety is the most significant risk faced by the Group's business. In this regard, the Group has always been focusing on building and perfecting a traceable platform, and adhering to the principle of 100–1 = 0. The Group lists food safety as one of the key tasks of the ESG team and establishes a three-level defense line for quality control so as to minimize the potential adverse impact of food safety incidents that may occur in the operation of the Group.

3. [*specific* condition] *Excerpt from Nissin Food's annual report (2023):*

Food safety risks

The Group's success depends on its branding and reputation and consumers' recognition and trust in its products. Food safety is the Group's lifeblood. Food safety incidents may damage its reputation and cause a loss of consumers' trust in its products. The profitability of the Group may be impacted as a result.

As a food manufacturer, the Group places high emphasis on food safety and quality management to ensure the health and safety of consumers. Therefore, the food production plants of the Group are all certified under ISO 22000 (an international standard for quality management systems), FSSC 22000 (an international standard that incorporates food defence approaches) reflecting the Group's commitment to a high standard of food safety, and HACCP system (except Nissin Vietnam) to identify hazardous substances and reduce risk on production. These certifications also ensure the production of safe and high-quality food products.

The Group has also applied rigorous analysis and inspections on raw materials and products. Food Safety Institute located in Shanghai, Mainland China, in which the Group owns a 5 % stake, provides technical support to the Group to perform quality inspections in each stage of production, from raw materials to final products. The Food Safety Institute is accredited by the China National Accreditation Service for Conformity Assessment ("CNAS") with ISO/IEC 17025. This shows that the Food Safety Institute's operating system and testing capabilities are nationally recognised. To further demonstrate the professionalism of the Food Safety Institute, it has obtained a certificate from the China Inspection Body and Laboratory Mandatory Approval ("CMA"). The CMA Certification proves that the Food Safety Institute has been recognised by the State.

Appendix B. RMD Manipulation in Experiment 1

(1) **Food Safety Risks**

Potential risks: As we cannot live without food, food safety should be a top priority. Once the Company experiences food safety problems due to management negligence, inadequate quality control system and other causes, it may face product liability claims, a negative brand impact or government penalties, which may adversely affect the Company's reputation and cause a loss of economic interests.

Risk management plans (generic)[10]**:** To respond to food safety risks, our company is optimizing its food safety testing methods, enhancing its multi-system accreditation and strengthening all aspects of food safety control.

Risk management plans (specific)**:** To respond to food safety risks, our company.

(1) is optimizing our food safety testing methods by:

① setting up a high-standard testing center equipped with top international testing facilities to conduct safety testing throughout the chain, from raw materials to finished products; and.

② conducting regular random quality inspections of products that have achieved market transformation—in 2021, over 1,400 batches of sampling inspection were implemented, and the pass rate was 100 %.

(2) enhancing our multi-system accreditation by:

① continuously passing qualification certifications such as ISO9001, ISO22000 and ISO14001; and.

② developing corporate standards for key processes that are more stringent than national standards.

(3) strengthening all aspects of food safety control by:

① improving our product quality and safety traceability system, such as newly introduced drinks with "one box, one code" to achieve full traceability; and

② strictly administering and supervising franchisees and stores and requiring our main suppliers to establish good quality control systems.

2. **Market Competition Risks**

Potential risks: The food industry's low barriers to entry and consumers' highly variable tastes lead to intense market competition. A failure of the Company to maintain continuous product innovation and stable product quality may result in decreases in consumer loyalty and purchase rates and erosion of the market share by competitors, thus affecting the Company's profitability.

Risk management plans (generic)**:** In response to market competition risks, our company understands changing trends in consumer behavior and is exploring new opportunities through multiple channels.

Risk management plans (specific)**:** In response to market competition risks, our company is

(1) understanding changing trends in consumer behavior by:

① using an e-commerce platform database to track product sales and conducting site visits to determine public opinions regarding the price, taste, packaging and other dimensions of our products, aiming to develop optimization plans for old products and R&D plans for new products; and.

② launching upgraded sugar-free snack foods and insisting on natural and additive-free to align with the health concept of reduced and low sugar among younger consumers.

(2) exploring new opportunities through multiple channels by:

① launching a brand pavilion in Tmall Supermarket and promoting activities through Weibo topics, Douyin short videos and B-station information flow; and

② dividing consumer groups into housewives, fitness enthusiasts, children and others and designing targeted marketing strategies based on the preferences of these groups, such as launching cookies for children in collaboration with animation IP.

[10] Contents in parentheses are not shown to the participants, the same below.

Appendix C. RMD Manipulation in Experiment 2

1. **Food Safety Risks**

Potential risks: As we cannot live without food, food safety should be a top priority. Once the Company experiences food safety problems due to management negligence, inadequate quality control system and other causes, it may face product liability claims, a negative brand impact or government penalties, which may adversely affect the Company's reputation and cause a loss of economic interests.

Risk management plans (specific)**:** To respond to food safety risks, our company is

(1) optimizing our food safety testing methods by:

① introducing advanced testing facilities and technology. In 2022, we invested up to USD 10 million for the continuous evaluation and improvement of our testing process on a quarterly basis.

② conducting regular random quality inspections of products that have achieved market transformation. In 2022, we implemented up to 1,400 batches of sampling inspection, with a passing rate of 100 %.

(2) strengthening all aspects of food safety control by:

① adopting strict food safety control measures at all stages of production, processing, storage and transportation. We clean and sterilize all aspects of the production workshop on a daily basis.

② improving our product quality and safety traceability system. We make every effort to use "one box, one code" to achieve full traceability.

Risk management plans (generic-listed)**:** To respond to food safety risks, our company is

(1) optimizing our food safety testing methods by:

① introducing advanced testing facilities and technology and evaluating and improving our testing processes on a regular basis; and.

② implementing advanced facilities and technology and conducting regular random quality inspections of products that have achieved market transformation to improve our product qualification rate and safety.

(2) strengthening all aspects of food safety control by:

① adopting strict food safety control measures at all stages of production, processing, storage and transportation and implementing prevention and control measures in response to food safety risks;

② improving our product quality and safety traceability system during production, processing, storage, transportation and other stages.

Risk management plans (generic-blocklike)**:** To respond to food safety risks, our company is

optimizing food safety testing methods by introducing advanced testing facilities and technology to evaluate and improve our testing processes on a regular basis; applying advanced facilities and technology and conducting regular random quality inspections of products that have achieved market transformation to improve our product qualification rate and safety.

strengthening all aspects of food safety control by adopting strict food safety control measures at all stages of production, processing, storage and transportation and implementing various prevention and control measures in response to food safety risks; improving our product quality and safety traceability system in the production, processing, storage, transportation and other stages.

2. **Market Competition Risks**

Potential risks: The food industry's low barriers to entry and consumers' highly variable tastes lead to intense market competition. A failure of the Company to maintain continuous product innovation and stable product quality may result in decreases in consumer loyalty and purchase rates and erosion of the market share by competitors, thus affecting the Company's profitability.

Risk management plans (specific)**:** To respond to market competition risks, our company is

(1) understanding changing trends in consumer behavior by:

utilizing the e-commerce platform database to track product sales, conducting quarterly comprehensive market surveys to understand consumers' needs and preferences and developing new products and promotional activities based on this information. In 2022, we launched 10 new products and not only retained our previous customers but also attracted up to 1 million new customers.

(2) exploring new opportunities through multiple channels by:

① making full use of Internet platforms. We have launched a brand pavilion in Tmall Supermarket and promoted activities through Weibo topics, Douyin short videos and bilibili information streaming.

② classifying consumers as, for example, housewives, fitness enthusiasts and children and designing targeted marketing strategies based on these groups' preferences. For example, we launched animation IP co-branded snacks for children and received a good market response.

Risk management plans (generic-listed)**:** To respond to market competition risks, our company is

(1) understanding changing trends in consumer behavior by:

utilizing an e-commerce platform database to track product sales, conducting quarterly comprehensive market surveys to understand consumers' needs and preferences and developing new products and promotional activities based on this information. We are actively investing in new products and promotional activities to increase the frequency of new product introductions, consolidate the stickiness of existing consumers and increase the number of new customers.

(2) exploring new opportunities through multiple channels by:

① exploiting new sales channels on e-commerce platforms and promoting activities through different types of social media to achieve our brand building goals;

② dividing consumers into different categories and designing targeted marketing strategies based on their preferences. We attract target consumer groups through innovative marketing methods and channels.

Risk management plans (generic-blocklike)**:** To respond to market competition risks, our company is

understanding the changing trends in consumer behavior by using an e-commerce platform database to track product sales, conducting quarterly comprehensive market surveys to understand consumers' needs and preferences and developing new products and promotional activities based on this information. We are actively investing in new products and promotional activities to increase the frequency of new product introductions, consolidate the stickiness of existing consumers and increase the number of new customers.

exploring new opportunities through multiple channels by exploiting new sales channels on e-commerce platforms and promoting activities through different types of social media to achieve our brand building goals and by dividing consumers into different categories and designing targeted marketing strategies based on their preferences. We attract target consumer groups through innovative marketing methods and channels.

References

Asay, H.S., Elliott, W.B., Rennekamp, K., 2017. Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *Account. Rev.* 92 (4), 1–25.

Abdullah, M., Abdul Shukor, Z., Mohamed, Z.M., Ahmad, A., 2015. Risk management disclosure: A study on the effect of voluntary risk management disclosure toward firm value. *J. Appl. Acc. Res.* 16 (3), 400–432. https://doi.org/10.1108/JAAR-10-2014-0106.

Aigbovo, O., Ilaboya, O., 2019. Does behavioural biases influences individual investment decisions. *Manag. Sci. Rev.* 10 (1), 68–89.

Arikan, O., 2021. The effect of boilerplate language on nonprofessional investors' judgments. *Account. Bus. Res.* 52 (4), 417–442. https://doi.org/10.1080/00014788.2021.1922990.

Barber, B.M., Odean, T., Zhu, N., 2008. Do Retail Trades Move Markets? *Rev. Financ. Stud.* 22 (1), 151–186. https://doi.org/10.1093/rfs/hhn035.

Beattie, V., McInnes, B., Fearnley, S., 2004. A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes. *Account. Forum* 28 (3), 205–236. https://doi.org/10.1016/j.accfor.2004.07.001.

Campbell, J., Chen, H., Dhaliwal, D., Lu, H., Steele, L., 2014. The information content of mandatory risk factor disclosures in corporate filings. *Rev. Acc. Stud.* 19 (1), 396–455. https://doi.org/10.1007/s11142-013-9258-3.

Cazier, R.A., McMullin, J.L., Treu, J.S., 2021. Are lengthy and boilerplate risk factor disclosures inadequate? An examination of judicial and regulatory assessments of risk factor language. *Account. Rev.* 96 (4), 131–155. https://doi.org/10.2308/TAR-2018-0657.

China Securities Regulatory Commission (CSRC), 2012. Guidelines for the Content and Format of Information Disclosure by Companies Publicly Issuing Securities No. 2 - Content and Format of Annual Reports. Available at: http://www.csrc.gov.cn/csrc/c101864/c1024633/content.shtml.

Elshandidy, T., Neri, L., 2015. Corporate governance, risk disclosure practices, and market liquidity: comparative evidence from the UK and Italy. *Corp. Gov.* 23 (4), 331–356.

Elshandidy, T., Zeng, C., 2022. The value relevance of risk-related disclosure: does the tone of disclosure matter? *Borsa Istanbul Rev.* 22 (3), 498–514. https://doi.org/10.1016/j.bir.2021.06.014.

Financial Reporting Council (FRC), 2018. Guidance on the Strategic Report. London: Financial Reporting Council. https://www.frc.org.uk/getattachment/fb05dd7b-c76c-424e-9daf-4293c9fa2d6a/Guidance-on-the-Strategic-Report-31-7-18.pdf.

He, Y., Tan, H., Yeo, F., Zhang, J., 2019. When do qualitative risk disclosures backfire? The effects of a mismatch in hedge disclosure formats on investors' judgments. *Contemp. Account. Res.* 36 (4), 2093–2112. https://doi.org/10.1111/1911-3846.12518.

Henry, E., Peytcheva, M., 2020. Joint effects of boilerplate and text markup on the judgments of novice and experienced users of financial information. *Behav. Res. Account.* 32, 1–20. https://doi.org/10.2308/bria-52582.

Hirst, D.E., Koonce, L., Venkataraman, S., 2007. How disaggregation enhances the credibility of management earnings forecasts. *J. Account. Res.* 45 (4), 811–837. https://doi.org/10.1111/j.1475-679X.2007.00252.x.

Hope, O., Hu, D., Lu, H., 2016. The benefits of specific risk-factor disclosures. *Rev. Acc. Stud.* 21 (4), 1005–1045. https://doi.org/10.1007/s11142-016-9371-1.

Huber, J., Palan, S., Zeisberger, S., 2019. Does investor risk perception drive asset prices in markets? Experimental evidence. *J. Bank. Financ.* 108, 105635. https://doi.org/10.1016/j.jbankfin.2019.105635.

Ibrahim, A.E.A., Aboud, A., 2023. Corporate risk disclosure and cost of capital: does measurement matter? *Int. J. Financ. Econ.*. https://doi.org/10.1002/ijfe.2862.

Jia, J., Li, Z., 2022. Risk management committees and readability of risk management disclosure. *J. Contemp. Account. Econ.* 18 (3), 100336. https://doi.org/10.1016/j.jcae.2022.100336.

Kahneman, D., Tversky, A., 1982. In: *The Simulation Heuristic. Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, New York, pp. 201–208.

Kasznik, R., Lev, B., 1995. To warn or not to warn: management disclosures in the face of an earnings surprise. *Account. Rev.* 70 (1), 113–134 http://www.jstor.org/stable/248391.

Koehler, D.K., 1991. Explanation, imagination, and confidence in judgment. *Psychol. Bull.* 110 (3), 499–519. https://doi.org/10.1037/0033-2909.110.3.499.

Koonce, L., Miller, J., Winchel, J., 2015. The effects of norms on investor reactions to derivative use. *Contemp. Account. Res.* 32 (4), 1529–1554. https://doi.org/10.1111/1911-3846.12118.

Kravet, T., Muslu, V., 2013. Textual risk disclosures and investors' risk perceptions. *Rev. Acc. Stud.* 18 (4), 1088–1122. https://doi.org/10.1007/s11142-013-9228-9.

Lang, M., Stice-Lawrence, L., 2015. Textual analysis and international financial reporting: large sample evidence. *J. Account. Econ.* 60 (2–3), 110–135. https://doi.org/10.1016/j.jacceco.2015.09.002.

Leitter, Z., Koonce, L., White, B.J., 2022. Business acquisition disclosures: the effect of identifying intangible assets on investors' judgments. Nanyang Bus. School Res. Paper 21–17. https://doi.org/10.2139/ssrn.3789990.

Levy, H., Shalev, R., Zur, E., 2018. The effect of CFO personal litigation risk on firms' disclosure and accounting choices. *Contemp. Account. Res.* 35 (1), 434–463. https://doi.org/10.1111/1911-3846.12378.

Li, Y., Wang, D., Meng, D., Hu, Y., 2024. Peer effect on climate risk information disclosure. *China J. Account. Res.* 17 (3), 100375.

Libby, R., Bloomfield, R., Nelson, M.W., 2002. Experimental research in financial accounting. *Acc. Organ. Soc.* 27 (8), 775–810. https://doi.org/10.1016/S0361-3682(01)00011-3.

Linsley, P.M., Lawrence, M.J., 2007. Risk reporting by the largest UK companies: readability and lack of obfuscation. *Account. Audit. Account. J.* 20 (4), 620–627. https://doi.org/10.1108/09513570710762601.

Liu, L., Geng, H., Wang, Y., 2025. How do managers use MD&A disclosures to respond to negative news? *China J. Account. Res.* 18 (1), 100405.

Nelson, K.K., Pritchard, A.C., 2016. Carrot or stick? The shift from voluntary to mandatory disclosure of risk factors. *J. Empir. Leg. Stud.* 13 (2), 266–297. https://doi.org/10.1111/jels.12115.

Rennekamp, K., 2012. Processing fluency and investors' reactions to disclosure readability. *J. Account. Res.* 50 (5), 1319–1354. https://doi.org/10.1111/j.1475-679X.2012.00460.x.

Rogers, J.L., Van Buskirk, A., Zechman, S.L., 2011. Disclosure tone and shareholder litigation. Account. Rev. 86 (6), 2155–2183. https://doi.org/10.2308/accr-10137.

Securities and Exchange Commission (SEC), 1998. A Plain English Handbook: How to Create Clear SEC Disclosure. SEC Office of Investor Education and Assistance. Retrieved from the SEC's. Website. http://www.sec.gov/pdf/handbook.pdf.

Securities and Exchange Commission (SEC), 2016. Business and Financial Disclosure Required by Regulation S-K. Release Nos. 33-10064, 34-77599. Available at: https://www.sec.gov/rules/concept/2016/33-10064.pdf.

Securities and Exchange Commission (SEC), 2019b. Modernization ofregulation S-K items 101, 103, and 105 (concept release nos. 33-100668; 34-86614. Available at: https://www.sec.gov/rules/proposed/2019/33-10668.pdf.

Sedor, L.M., 2002. An explanation for unintentional optimism in analysts' earnings forecasts. *Account. Rev.* 77 (4), 731–753. https://doi.org/10.2308/accr.2002.77.4.731.

Schoemaker, P., 1991. When and how to use scenario planning: a heuristic approach with illustration. *J. Forecast.* 10, 549–564. https://doi.org/10.1002/for.3980100602.

Shearman & Sterling LLP, 2014. https://www.sec.gov/comments/disclosure-effectiveness/disclosureeffectiveness-25.pdf.

Tan, H.T., Wang, E.Y., Zhou, B., 2014. When the use of positive language backfires: the joint effect of tone, readability, and investor sophistication on earnings judgments. *J. Account. Res.* 52 (1), 273–302.

Tan, H.T., Yeo, F., 2023. You have been forewarned! The effects of risk management disclosures and disclosure tone on investors' judgments. *Acc. Organ. Soc.* 105, 101400. https://doi.org/10.1016/j.aos.2022.101400.

Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. *Cogn. Psychol.* 5 (2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9.

Wang, K., 2021. Is the tone of risk disclosures in MD&As relevant to debt markets? Evidence from the pricing of credit default swaps. *Contemp. Account. Res.* 38 (2), 1465–1501. https://doi.org/10.1111/1911-3846.12644.

Yao, Y., Wei, L., Jing, H., Chen, M., Li, Z., 2024. The impact of readability of risk disclosures in bond prospectuses on credit risk premium. *Res. Int. Bus. Financ.* 70, 102310. https://doi.org/10.1016/j.ribaf.2024.102310.

Zhao, W., Yang, H., Zhou, H., 2022. Linguistic specificity and stock price synchronicity. *China J. Account. Res.* 15 (1), 100219.

Zukier, H., 1982. The dilution effect: The role of the correlation and the dispersion of predictor variables in the use of nondiagnostic information. J. Personality Soc. Psychol. 43 (6), 1163 https://psycnet.apa.org/doi/10.1037/0022-3514.43.6.1163.

China Journal of Accounting Research 18 (2025) 100435

HOSTED BY

Contents lists available at ScienceDirect

China Journal of Accounting Research

journal homepage: www.elsevier.com/locate/cjar

Acquaintance means booster? Why stable customers matter for firm productivity

Yunjiang Geng, Yuelong Wang*, Shuaijun Wu

School of Accounting, Dongbei University of Finance and Economics, Dalian 116025 Liaoning, China

ARTICLE INFO

Article history:
Received 3 December 2024
Accepted 2 July 2025
Available online 16 July 2025

Keywords:
Customer stability
Total factor productivity
Type I agency cost
Firm reputation
Supply chain spillover effect

ABSTRACT

Amid sluggish global growth and rising uncertainties, supply chain stability is vital for sustaining economic production. Although studies examine the impacts of supply chain relationships on firm performance, their effect on total factor productivity (TFP) remains unexplored. Using data from 1559 A-share listed companies in China (2008–2022), this study examines customer stability's impact on TFP and finds that customer stability enhances TFP by reducing Type I agency costs and improving firm reputation. It also generates significant spillover effects, increasing customer TFP through supply chain finance. This effect is more pronounced for firms in high-tech industries and regions with higher marketization and social trust. These findings offer new insights into enhancing firms' efficiency through effective supply chain relationship management.

1. Introduction

In the context of the current global economic stagnation and increasing external uncertainties, enhancing total factor productivity (TFP) has emerged as a critical pathway for countries to achieve high-quality economic development (Bellocchi et al., 2021). TFP refers to the "residual" portion of total output that cannot be explained by input factors and is a comprehensive indicator of industrial upgrading efficiency and overall economic performance (Baier et al., 2006). Micro-enterprises form the bedrock of the macroeconomy, and the improvement of macroeconomic output efficiency ultimately depends on increasing TFP within these firms. Therefore, enhancing firm-level TFP is pivotal to steering the macroeconomy toward high-quality and sustainable development.

* Corresponding author.
E-mail addresses: gyjiang@dufe.edu.cn (Y. Geng), wangyuelonglb@163.com (Y. Wang).

https://doi.org/10.1016/j.cjar.2025.100435
1755-3091/

Concurrently, with globalization and the integration of the world economy, ensuring the security and stability of industrial and supply chains has become crucial for safeguarding national economic production. From a microeconomic perspective, supply chain security and stability are critical for ensuring steady firm-level production. As one of the most crucial characteristics of supply chain transactions, customer stability is widely used to evaluate the safety and stability of supply chains. Customer stability reflects the long-term relationships between focal firms and their customers, and such enduring relationships are a significant resource advantage for firms (Bauer et al., 2018). Abnormal fluctuations in these relationships may attract third-party scrutiny. For instance, in 2022, Qingyun Technology (688316.SH), a listed company in China, was required by the Shanghai Stock Exchange to explain the frequent changes in its major customers, highlighting regulatory concerns about potential risks associated with customer volatility. Against this backdrop, more firms are inclined to establish long-term and stable partnerships with trade partners (Patatoukas, 2012; Hui et al., 2019). This has fueled significant academic interest in the economic consequences of maintaining customer stability. Evidence suggests that stable customer relationships can reduce customer acquisition costs (Baiman and Rajan, 2002), improve market share (Patatoukas, 2012) and enhance innovation performance (Bernard et al., 2019). Moreover, stable customer relationships facilitate access to trade credit (Zhang et al., 2024), mitigate financial constraints (Kim et al., 2015) and reduce operational risks (Fornell et al., 2006).

Although numerous studies discuss the benefits of customer stability, none have specifically examined its impact on a firm's TFP. Due to inefficiencies resulting from resource misallocation, Chinese firms have low TFP, indicating considerable potential for improvement (Zhang et al., 2023). Moreover, although China is the largest manufacturing country with the most comprehensive industrial system and the largest developing economy, the stability and flexibility of its industrial and supply chains still require further strengthening. Therefore, investigating whether customer stability within the supply chains of Chinese firms can enhance TFP holds significant practical relevance. Furthermore, Serpa and Krishnan (2018) find that the characteristics of other firms within the supply chain can affect a firm's TFP. Given that customer stability reflects the relationship dynamics between suppliers and customers, whether the customer stability of focal firms influences their customers' TFP remains to be determined. This potential spillover effect within the supply chain warrants further investigation.

Based on the aforementioned analysis, this study addresses the following four questions. First, can stable customer relationships effectively enhance a firm's TFP? Second, if so, what is the primary mechanism underlying this effect? Third, does this impact propagate along the supply chain, creating a spillover effect? Fourth, how does the estimated effect vary across different firms?

To address the first two research questions, we construct our primary dataset by selecting A-share listed companies in China from 2008 to 2022 using the China Stock Market and Accounting Research (CSMAR) database, resulting in a final sample of 1559 firms with 6005 firm-year observations. Our empirical analysis demonstrates that customer stability significantly enhances a firm's TFP. To mitigate potential concerns regarding sample selection bias, endogeneity and reverse causality, we perform multiple robustness tests, including the Heckman two-stage method, propensity score matching (PSM), staggered difference-in-differences (DID) analysis and an instrumental variable approach, all of which consistently support our main conclusion. Furthermore, mediation analysis indicates that customer stability primarily enhances TFP through two channels: reducing Type I agency costs and bolstering firm reputation.

To further examine whether customer stability generates a TFP spillover effect within the supply chain, we manually construct a unique dataset comprising 750 matched "focal firm–customer–year" dyadic observations. The results reveal that focal firms' customer stability also improves their downstream customers' TFP, indicating a significant supply chain spillover effect. By supplementing our analysis with inter-firm trade credit data and employing three identification strategies, we provide compelling evidence that these spillover effects mainly stem from the supply chain financial support extended by focal firms to their customers. This finding indicates that customer stability not only enhances a firm's own productivity but also generates positive externalities through the supply chain finance mechanism.

Our heterogeneity analysis further reveals important contextual variations in the productivity effect of customer stability. The results of subgroup regression show that the TFP-enhancing effect of customer stability is particularly pronounced among high-tech firms and firms located in regions with higher levels of marketization and social trust. These findings not only reinforce our main conclusion but also offer valuable implications for firm management practices and policymaking.

Compared with the literature, this study makes several key contributions. First, while previous studies examine factors influencing TFP from a supply chain perspective (Gu et al., 2023; Xu and Guan, 2023), the exploration of how supply chain relationships affect TFP remains limited. This study focuses on the impact of customer stability, a critical characteristic of supply chain relationships, on firm TFP, thereby expanding the scope of TFP research. Second, although the significant influence of stable supply chain relationships on operational risk (Fornell et al., 2006), firm revenue (Gosman and Kohlbeck, 2009), supply chain finance (Liu et al., 2022a) and trade credit (Zhang et al., 2024) is well documented, their impact on firm productivity has been overlooked. This study fills that gap by demonstrating that customer stability enhances firm TFP through governance and reputation effects, thereby clarifying the mechanisms underlying supply chain customer management and contributing to supply chain management theory while also expanding the literature on the economic outcomes of customer stability. Third, previous research shows that customers and suppliers in vertical supply chain relationships can influence each other's productivity (Serpa and Krishnan, 2018). Moreover, factors such as ESG (Tang et al., 2023) and digital transformation (Li et al., 2024) generate supply chain spillover effects. This study finds that customer stability, as a relational characteristic between suppliers and customers, also exhibits a significant spillover effect. Specifically, greater customer stability not only enhances the TFP of focal firms but also improves that of their customers, thereby contributing to the broader literature on supply chain spillover effects.

The remainder of this study is structured as follows. Section 2 reviews the relevant literature and develops the research hypotheses. Section 3 describes the sample selection process, variable definitions and empirical model. Section 4 presents the main empirical results, including robustness checks and mechanism analyses. Section 5 offers additional analyses. Section 6 concludes the study and discusses its implications.

2. Literature review and hypothesis development

2.1. Customer stability and TFP

Customer stability, a key characteristic of supply chain relationships, refers to repeated trading activities between customers and suppliers (Yang, 2017). It functions as an informal institution that provides external governance, particularly when major customers are involved because they alone have the capacity and motivation to influence the firm (Serpa and Krishnan, 2018). Studies often measure customer stability based on the stability of the largest customer (Li and Yang, 2011). The presence of major customers sends positive signals to the market (Hui et al., 2019), and firms tend to invest substantial proprietary resources to maintain relationships with these customers to achieve "relational transactions." However, some major customers may erode firm profits and reduce earnings quality (Gosman et al., 2004). Excessive dependence on major customers may increase operational risks. In response, firms may engage in upward earnings management to mitigate these risks (Healy et al., 1999), thereby affecting their informational environment.

The aforementioned studies suggest that firms are particularly susceptible to the influence of major customers (Chen et al., 2022). In practice, an increasing number of firms prefer to establish long-term partnerships with a limited number of trading partners (Patatoukas, 2012; Hui et al., 2019). Thus, academic research increasingly focuses on the economic implications of customer stability. In terms of cost performance and risk, stable customer relationships reduce the resources and costs associated with acquiring new customers (Baiman and Rajan, 2002), leverage supply chain information advantages and lower communication and monitoring expenses. These relationships also enable firms to develop flexible cost-adjustment mechanisms that mitigate operational risks (Fornell et al., 2006). Regarding resource allocation and innovation, stable customer relationships, viewed as intangible assets formed through "organizational capital," are recognized by capital markets and reflected in firm valuation (Gosman et al., 2004). These relationships help alleviate financial constraints (Kim et al., 2015), enhance access to external financing and reduce borrowing costs (Petersen and Rajan, 1997; Hertzel et al., 2008). They also reduce the burden of trade credit and working capital requirements from customers (Hofmann and Kotzab, 2010), optimize cash holdings and provide more resources for research and development (R&D) and innovation, ultimately improving investment efficiency and innovation performance (Bernard et al., 2019). In the context of the firm's informational environment, the presence of stable customers provides valuable insights to both internal and external stakeholders (Bonacchi et al., 2015), attracts increased analyst attention and enhances information transparency while reducing information asymmetry (Wang and Peng, 2016).

Another key focus of this study is firm-level TFP, which captures the effects of technological progress, managerial efficiency and the institutional environment on output (Cheng and Zhu, 2024). Technological advancements and improvements in resource allocation efficiency are key drivers of TFP growth (Hu et al., 2015). However, scholars also identify information asymmetry (Guo and Zhang, 2023) and financial constraints (Hopenhayn, 2014) as significant barriers to TFP improvement. According to stakeholder theory, a firm's production and operations are closely connected with its stakeholders, including employees, institutional investors, customers, suppliers, communities, government agencies, industry associations and unions. For example, highly skilled employees (Wang et al., 2023), government-provided tax relief (Liu et al., 2022b) and higher productivity among supply chain partners (Serpa and Krishnan, 2018) all enhance firm TFP through various channels.

In summary, although many scholars highlight the economic benefits of maintaining stable customer relationships, direct evidence on whether customer stability improves a firm's TFP remains limited, particularly given the growing importance of securing industrial and supply chains and promoting high-quality firm development.

2.2. Impact of customer stability on a firm's TFP

Resource-based theory (Wernerfelt, 1984) posits that a firm comprises a portfolio of diverse resources, with the scarcity and strategic allocation of these resources determining its competitive advantage. Stable customer relationships, regarded as critical economic resources developed through long-term and frequent transactions (Möller and Halinen, 1999), positively influence a firm's production and operational decisions. Um and Kim (2019) find that robust supply chain collaboration not only effectively reduces transaction costs among firms within the chain but also enhances firm performance. Similarly, Woo and Suresh (2022) examine the involvement of Korean firms in supply chain cooperation policies and report that stable supply chain partnerships enhance firm performance, specifically in terms of return on assets, return on investment and the efficiencies achieved in marketing, production and inventory management.

All this evidence collectively indicates that stable supply chain relationships are crucial sources of enhanced operational performance and production efficiency. Thus, this study focuses on customer stability, a specific characteristic of customer relationships, and examines its potential correlation with a firm's TFP.

First, stable customers can significantly reduce transaction costs and operational risks, thereby improving TFP. The inherent incompleteness of formal mechanisms, such as market regulation and transactional contracts, often leads to high transaction costs and resource inefficiencies (Saussier, 2000). However, stable customer relationships mitigate these shortcomings by facilitating "relational transactions" within internal networks (Allen et al., 2005). Maintaining existing customers is substantially less costly than acquiring new ones (Baiman and Rajan, 2002). Furthermore, stable customers help ensure the fulfillment of sales contracts, reduce the likelihood of customer defaults and lower the firm's overall risk exposure (Fornell et al., 2006). They also contribute to consistent growth in sales and market share, providing the firm with more stable revenue streams (Gosman and Kohlbeck, 2009) and thus enhancing its risk-bearing capacity.

Second, stable customers can enhance a firm's financing capabilities and innovation efficiency, thereby providing intrinsic momentum for improving TFP. Financing constraints and technological innovation capabilities are key factors influencing TFP (Hopenhayn, 2014; Hu et al., 2015). Supply chain finance, grounded in stable cooperative relationships, provides firms with greater access to trade credit and helps alleviate financing constraints (Hertzel et al., 2008; Kim et al., 2015). Thus, firms obtain more resources for investment and R&D activities, which improve investment efficiency and innovation performance (Bernard et al., 2019). Additionally, collaboration among upstream and downstream supply chain partners facilitates technological advancements and fosters collaborative innovation performance (Shen et al., 2021).

Finally, stable customers can enhance a firm's information transparency and provide incremental information to external markets, thereby fostering a favorable information environment and external oversight for improving TFP (Bonacchi et al., 2015). Higher customer stability signals consistent firm performance in both product and capital markets, attracts increased analyst coverage and improves earnings forecast quality (Wang and Peng, 2016). Moreover, increased customer stability promotes trust and information sharing within the supply chain (Panahifar et al., 2018), which not only enhances firm transparency but also strengthens external supervision and scrutiny.

In summary, the presence of stable customers reduces transaction costs and operational risks, increases financing capacity and innovation efficiency and cultivates a more transparent external environment, thereby creating favorable internal and external conditions for enhancing TFP. Based on the aforementioned considerations, we propose the following hypothesis:

H1. The higher a firm's customer stability is, the more it facilitates the enhancement of TFP.

2.3. Mediating role of Type I agency costs

Information asymmetry is a fundamental driver of agency costs. The separation of ownership and management results in Type I agency costs, which arise between shareholders and management. These agency costs lead to ineffective corporate governance and insufficient supervisory mechanisms, potentially impeding improvements in TFP (Chiang and Lin, 2007). Managerial self-interest, a manifestation of agency costs, directly reflects weak corporate governance, increases the likelihood of supply chain disruptions and reduces production efficiency (Bauer et al., 2018). In summary, the existence of Type I agency costs prompts management to prioritize personal interests over those of shareholders, thereby reducing operational efficiency and hindering TFP enhancement.

However, stable customers can mitigate these agency problems in three specific ways. First, customer stability directly affects firm profitability. Stable customers ensure steady sales growth (Gosman and Kohlbeck, 2009), which alleviates performance pressure on management and reduces self-serving behaviors such as short-termism, concealing operational failures and over-investment. Second, stable customers maintain close, long-term relationships with firms, thereby increasing external scrutiny and oversight from key customers. This pressure reduces managerial opportunism and self-interested behavior by increasing the risk of potential supply chain disruptions. Finally, greater customer stability promotes information sharing and collaboration within the supply chain (Panahifar et al., 2018), effectively reducing information asymmetry (Wang and Peng, 2016) and Type I agency costs. Based on this analysis, we propose the following hypothesis:

H2. Customer stability enhances a firm's TFP by reducing Type I agency costs.

2.4. Mediating role of firm reputation

Firm reputation represents a comprehensive assessment of a firm's past conduct, transmitting multidimensional signals, including the quality of accounting information, to external stakeholders, thereby serving as an efficient screening mechanism. A strong reputation is a valuable intangible asset that directly supports a firm's long-term sustainability (Freeman et al., 2007) and significantly affects production efficiency (Stuebs and Sun, 2010). A positive reputation creates a "snowball effect" that offers several advantages, such as reducing information asymmetry (Guan and Zhang, 2019), increasing sales revenue and improving capital market performance. Moreover, a positive reputation helps constrain opportunistic behavior in contractual relationships, thereby enhancing production efficiency (Chen et al., 2022). Thus, a strong reputation is a critical resource for maintaining competitive advantage and improving TFP.

Simultaneously, stable customer relationships provide incremental information that signals operational stability, product or service reliability and strong risk-bearing capacity, thereby enhancing firm reputation. In China, customer information is disclosed voluntarily; firms that proactively disclose stable customer data attract greater analyst coverage and improve the quality of earnings forecasts (Wang and Peng, 2016). Furthermore, higher customer stability reflects long-term relationships that foster trust among external stakeholders. These enduring partnerships ensure consistent revenue streams (Gosman and Kohlbeck, 2009), align with stakeholder expectations and further strengthen firm reputation. Based on the aforementioned analysis, we propose the third hypothesis:

H3. Customer stability enhances a firm's TFP by improving firm reputation.

The main conceptual model in this study is depicted in Fig. 1.

3. Methodology

3.1. Data and sample

This study selects Chinese A-share firms listed on the Shanghai and Shenzhen stock exchanges from 2008 to 2022 as the initial sample. In 2007, the China Securities Regulatory Commission (CSRC) issued the "Standards on the Contents and Formats of Information Disclosure by Companies Offering Securities to the Public No. 2—Contents and Formats of Annual Reports (2007 Revision)," which stipulates that "companies should disclose information about their major suppliers and customers: the proportion of the total annual procurement accounted for by purchases from the top five suppliers, and the proportion of the total annual sales accounted for by sales to the top five customers." Thus, from 2008 onward, listed firms began formally and extensively disclosing information about their top five customers in their annual reports. Accordingly, this study designates 2008 as the starting year for the sample. However, because base-year data are required to calculate key independent variables, the final dataset spans from 2009 to 2022, covering a total of 14 years. Table 1 shows the sample distribution by year and industry.

In accordance with the research requirements, the initial sample is filtered by (1) excluding financial industry firms based on the "Guidelines for Industry Classification of Listed Companies (2012 Revision)" issued by the CSRC; (2) manually excluding samples that did not disclose customer names in detail; (3) excluding firms listed for less than 2 years because measuring the core independent variable, customer stability, requires 2 consecutive years of annual report data; (4) excluding firms designated with special treatment (ST and ST*); (5) excluding observations of firms that delisted during the sample period; and (6) eliminating observations with missing values for the variables of interest. This process results in a final sample of 6005 observations. To mitigate the influence of extreme values, all continuous variables are winsorized at the 1 % and 99 % levels. The financial data of listed companies used in this study are obtained from the CSMAR database, and customer information is obtained from the Chinese Research Data Services database.

3.2. Measures

3.2.1. Dependent variable

The estimation of a firm's TFP (*TFP*) remains a subject of considerable debate in the literature, with most scholars adopting the semi-parametric methods proposed by Olley and Pakes (OP; 1996) and Levinsohn and Petrin (LP; 2003). Because the OP method requires positive investment values, leading to the exclusion of many firm samples, this study adopts the LP method to measure firms' TFP in the baseline regression. For robustness checks, this study also employs the ordinary least squares (OLS) and fixed effects (FE) methods.

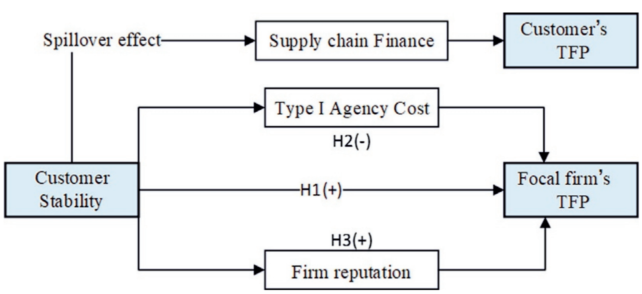


Fig. 1. Conceptual model.

Table 1
Final sample distribution.

Panel A: The distribution of sample firms by year

By year		Frequency (n)	Percentage (%)
	2009	88	1.47
	2010	565	9.41
	2011	755	12.57
	2012	868	14.45
	2013	873	14.54
	2014	434	7.23
	2015	330	5.50
	2016	358	5.96
	2017	353	5.88
	2018	342	5.70
	2019	304	5.06
	2020	285	4.75
	2021	256	4.26
	2022	194	3.23
Total		6005	100

Panel B: The distribution of sample firms by industry

Industries	Frequency (n)	Percentage (%)
Agriculture, hunting, forestry, fishing	73	1.22
Mining	244	4.06
Manufacturing	3543	58.99
Electricity, heat, gas, water production and supply	466	7.76
Construction	215	3.58
Wholesale and retail trade	273	4.55
Transportation, warehousing, and postal services	171	2.85
Hotel and catering sectors	20	0.33
Information transmission, software, and information technology services	389	6.48
Real estate	177	2.95
Leasing and business service	65	1.08
Scientific research and technical services	94	1.57
Water, environmental, and public utility management	102	1.70
Residential services, repairs, and other services	1	0.02
Education	3	0.05
Health and social work	2	0.03
Culture, sports, and entertainment	74	1.23
Other	93	1.55
Total	6005	100

By using multiple approaches, this study addresses potential endogeneity concerns and enhances the precision of TFP measurement.

3.2.2. Independent variable

The independent variable in this study is customer stability (*CS*). If a firm's top five customers from the previous year remain among its top five customers in the current year, they are classified as stable customers. Following Wang and Peng (2016) and Yang (2017), we define customer stability as the ratio of the number of stable customers over 2 consecutive years to 5. This ratio captures the strength of customers' repeated behaviors (Liu et al., 2022a) and effectively reflects the focal firm's customer stability. Additionally, the following variables are used for robustness checks: (1) a dummy variable that equals 1 if customer stability exceeds the sample mean and 0 otherwise (*CS_dummy*); (2) the natural logarithm of the number of stable customers plus 1 (*LnCS*); and (3) the ratio of the number of stable customers that consistently appear in the top five over the past 3 years to 15 (*CS_3*).

3.2.3. Mediating variables

We test two mediating variables, namely Type I agency costs and firm reputation, to elucidate the transmission mechanisms linking customer stability to TFP. Following Lei et al. (2023), we measure Type I agency costs (AC) using the widely adopted management expense ratio, defined as the ratio of management expenses to total operating income. A higher (lower) ratio indicates more (less) severe Type I agency cost problems.

Most studies use external ranking data or survey responses to assess firm reputation, although such methods are often subjective and one-dimensional. Following Guan and Zhang (2019), we comprehensively consider the evaluation of firm reputation by various stakeholders. Specifically, we select 12 firm reputation indicators: (1) rankings of firm assets, income, net profit and firm value within the industry from the perspective of consumers and society; (2) asset-liability ratio, current ratio and long-term debt ratio from the perspective of creditors; (3) earnings per share, dividend per share and whether the firm is audited by a Big Four accounting firm from the perspective of shareholders; and (4) sustainable growth rate and the proportion of independent directors from the perspective of the firm itself. We then perform factor analysis to calculate the firm reputation score. Based on these scores, we categorize firms into 10 groups, with each group assigned a value ranging from 1 to 10. A higher score indicates better firm reputation (*Rep*).

3.2.4. Control variables

In this study, we incorporate control variables that may affect firm TFP at various levels to mitigate endogeneity issues caused by omitted variables. First, we control for firm characteristics, including firm size (*Size*), firm age (*Age*) and firm ownership structure (*Soe*). *Size* is measured as the natural logarithm of total assets, and *Age* is measured as the natural logarithm of the number of years since the firm's establishment. *Soe* is a dummy variable that equals 1 if the firm is a state-owned enterprise and 0 otherwise. Second, given that financial status is also an important factor affecting firms' TFP, we control for the firms' financial accounting variables (Wang and Zhang, 2024). Specifically, we control for profitability (*Roa*), growth (*Growth*), capital structure (*Lev*), Tobin's Q (*TobinQ*) and the fixed asset ratio (*Fixed*). *Roa* is the ratio of net profit to total assets. *Growth* reflects the annual growth rate of primary business income. *Lev* is defined as the ratio of total debt to total assets. *TobinQ* is the ratio of market value to the replacement cost of assets. *Fixed* represents the ratio of fixed assets to total assets. Moreover, we control for firms' internal governance factors (Cheng and Zhu, 2024), including equity concentration (*Top1*), duality of the chairman and CEO (*Dual*), board size (*Board*) and the proportion of independent board members (*Indep*). *Top1* refers to the shareholding ratio of the largest shareholder. *Dual* is a dummy variable that equals 1 if the firm's chairman and CEO positions are held by the same individual and 0 otherwise. *Board* is measured as the natural logarithm of the number of board members. *Indep* represents the proportion of independent directors on the board. Detailed definitions of all variables are provided in Appendix A.

3.3. The empirical model

Following Guo and Zhang (2023) and Lin and Deng (2024), we adopt the following multivariate OLS regression model to examine the impact of customer stability on TFP:

$$TFP_{i,t} = \beta_0 + \beta_1 CS_{i,t} + \beta_2 \sum Controls_{i,t} + Ind + Year + Area + \varepsilon_{i,t} \tag{1}$$

All variables are described in Section 3.2. $Controls_{i,t}$ represents the control variables. The firm and year are described by the subscripts i and t, respectively. In addition, we control for the unobserved industry fixed effects, year fixed effects and area fixed effects. $\varepsilon_{i,t}$ is the error term. To alleviate the impact of firm-level clustering on standard errors, all regression results of our study are subjected to cluster adjustments at the firm level.

4. Analysis and results

4.1. Descriptive statistics and correlations

The results of the descriptive statistics for the aforementioned variables are shown in Table 2. For the dependent variable, the mean and median values of *TFP* are 8.165 and 8.076, respectively. The median is lower than the mean, indicating that the *TFP* of listed companies follows a normal right-skewed distribution and most of the listed companies have lower than average *TFP*. This finding is consistent with that reported by Wang and Zhang (2024). Moreover, *TFP* values range from 5.914 to 10.878, indicating a significant disparity in productivity among firms.

Table 3 presents the correlation coefficients among variables. The correlation coefficient between *CS* and *TFP* is 0.054 and is significant at the 1 % level, indicating a significant and positive correlation between *CS* and *TFP*. Several control variables, including *Size*, *Roa* and *Board*, also exhibit significant and positive correlations with *TFP*. This finding indicates the importance of controlling for these variables in our model. Moreover, none of the models in this study exhibit variance inflation factor values exceeding 3 (Table 4), suggesting that multicollinearity is not a concern in our analysis.

4.2. Overall hypothesis testing

This study estimates the impact of customer stability on firm TFP based on Eq. (1). The baseline regression results are presented in Table 5. Column (1) omits control variables, Column (2) includes all previously discussed control variables and Column (3) further incorporates year, industry and area fixed effects. The results in Table 5 show that the coefficient on *CS* is significant and positive across all specifications, indicating a notable positive impact of *CS* on *TFP*. Specifically, in Column (3), the coefficient of *CS* is positive and significant at the 1 % level (0.148, $t = 3.93$), supporting H1.

Moreover, the economic significance of improving customer stability is substantial. On average, a one standard deviation increase in *CS* (0.2643) results in a 0.039 (0.148×0.2643) increase in *TFP*, which accounts for 0.478 % (0.039/8.1651) of the average *TFP*. In other words, *TFP* could increase by 0.478 %. The average annual geometric growth rate of firm-level *TFP* measured using the LP method is only 0.479 % from 2009 to 2022. Therefore, improving customer stability has considerable economic significance for enhancing a firm's TFP.

Table 2
Descriptive statistics.

Variable	N	Mean	SD	Min	Median	Max
TFP	6005	8.1651	1.0254	5.9139	8.0756	10.8775
Stable	6005	0.4927	0.2643	0.0000	0.4000	1.0000
Size	6005	22.0886	1.2607	19.7263	21.9258	25.9096
Lev	6005	0.4465	0.2139	0.0508	0.4465	0.9107
Soe	6005	0.4828	0.4997	0.0000	0.0000	1.0000
Fixed	6005	0.2413	0.1750	0.0032	0.2059	0.7445
TobinQ	6005	1.9429	1.2318	0.8572	1.5313	8.0557
Growth	6005	0.1730	0.3775	−0.5284	0.1122	2.2531
Age	6005	2.8258	0.3610	1.6094	2.8904	3.4965
Indep	6005	36.9241	5.0182	30.7700	33.3300	57.1400
Dual	6005	0.2143	0.4104	0.0000	0.0000	1.0000
Top1	6005	35.4920	15.0986	8.4484	32.9887	74.9648
Board	6005	2.1646	0.1981	1.6094	2.1972	2.7081
Roa	6005	0.0554	0.0588	−0.1791	0.0523	0.2343

Table 3
Correlation coefficient.

	CS	*TFP*	*Size*	*Lev*	*Soe*	*Fixed*	*TobinQ*
CS	1						
TFP	0.054***	1					
Size	0.083***	0.750***	1				
Lev	−0.037***	0.465***	0.496***	1			
Soe	0.073***	0.261***	0.377***	0.344***	1		
Fixed	0.202***	−0.082***	0.190***	0.163***	0.295***	1	
TobinQ	−0.018	−0.320***	−0.454***	−0.208***	−0.145***	−0.113***	1
Growth	−0.132***	0.106***	0.011	0.023*	−0.073***	−0.055***	0.059***
Age	0.057***	0.129***	0.219***	0.211***	0.170***	0.029**	0.016
Indep	−0.049***	−0.028**	−0.024*	0.010	−0.097***	−0.053***	0.044***
Dual	−0.020	−0.152***	−0.191***	−0.149***	−0.301***	−0.125***	0.076***
Top1	0.058***	0.241***	0.278***	0.069***	0.257***	0.162***	−0.156***
Board	0.039***	0.179***	0.269***	0.144***	0.272***	0.181***	−0.134***
Roa	0.086***	0.156***	0.030**	−0.251***	−0.074***	−0.037***	0.125***
	Growth	*Age*	*Indep*	*Dual*	*Top1*	*Board*	*Roa*
Growth	1						
Age	−0.063***	1					
Indep	−0.004	−0.005	1				
Dual	0.017	−0.072***	0.122***	1			
Top1	0.005	−0.143***	0.019	−0.093***	1		
Board	−0.027**	−0.009	−0.481***	−0.157***	0.047***	1	
Roa	0.279***	−0.130***	−0.044***	0.007	0.110***	0.047***	1

Notes: $N = 6005$. ***, **, and * refer to significance at the 1 %, 5 %, and 10 % level, respectively. (two tailed test).

Table 4
Variance inflation factor test.

Variable	*Size*	*Lev*	*Soe*	*Fixed*	*TobinQ*	*Growth*	*Age*	*Indep*	*Dual*	*Top1*	*Board*	*Roa*	Mean
VIF	2.71	1.93	1.72	1.83	1.74	1.21	1.84	1.45	1.12	1.36	1.67	1.43	1.67

4.3. Robustness tests

4.3.1. Heckman two-stage method

Because the CSRC does not mandate listed companies to disclose detailed information on their top five customers, the customer stability data used in this study rely on voluntary disclosures. Firms with more stable customer relationships may be more inclined to disclose such information, thereby signaling positive image to the public. This tendency may introduce sample selection bias. To address this issue, following Wang and Peng (2016), we employ the Heckman two-stage method to mitigate endogeneity concerns related to self-selection bias.

In the first-stage regression, the disclosure of customer information is defined as a binary variable. Control variables, including *Size*, *Lev*, *Roa*, *Top1*, *Soe*, *Indep* and the degree of industry competition (measured using the Herfindahl–Hirschman Index), are included as exclusionary constraint variables. The inverse Mills ratio (*IMR*), derived from the first-stage regression, is then included as a control variable in the second-stage regression (Model 1). The results, presented in Column (1) of Table 6, indicate that the coefficient of *IMR* is significant at the 5 % level, confirming the presence of endogeneity in the original regression. Nevertheless, the coefficient of *CS* remains positive and significant (0.148, t value = 3.57), indicating that the main conclusion of this study remains robust after accounting for potential sample selection bias.

Table 5
Baseline regression results.

Variables	(1)	(2)	(3)
	TFP	*TFP*	*TFP*
Stable	0.208**	0.183***	0.148***
	(2.40)	(4.05)	(3.93)
Size		0.543***	0.558***
		(31.15)	(35.36)
Lev		1.021***	0.890***
		(10.72)	(10.16)
Soe		0.016	0.074**
		(0.40)	(2.07)
Fixed		−1.512***	−1.295***
		(−13.62)	(−12.14)
TobinQ		−0.010	−0.002
		(−0.93)	(−0.1864)
Growth		0.112***	0.126***
		(4.94)	(5.8060)
Age		−0.084*	−0.053
		(−1.82)	(−1.20)
Indep		−0.004	0.001
		(−1.30)	(0.15)
Dual		−0.048	−0.074***
		(−1.57)	(−2.84)
Top1		0.003***	0.003***
		(3.21)	(3.50)
Board		−0.0698	0.0444
		(−0.66)	(0.51)
Roa		2.721***	2.689***
		(12.49)	(13.63)
_ *Constant*	8.062***	−3.720***	−4.740***
	(153.32)	(−8.55)	(−9.20)
Year	No	No	Yes
Ind	No	No	Yes
Province	No	No	Yes
Observations	6005	6005	6005
Adj-R^2	0.003	0.667	0.754

Note: ***, **, and * refer to significance at the 1%, 5%, and 10% level, respectively. Unless otherwise specified, standard errors in the results below are clustered at the firm level.

4.3.2. PSM method

TFP is influenced by multiple factors, and the conclusions drawn thus far may be subject to omitted variable bias. To further mitigate the issue of selection bias—where firm-level characteristics may lead to variations in TFP—we apply the PSM method for validation. First, we divide the sample into two groups, high and low customer stability, based on the mean value of *CS*. Then, using the control variables from Model (1), we apply the 1:1 nearest neighbor matching method to match each high-*CS* sample with a low-*CS* sample that has the closest propensity score. Finally, we reanalyze the matched sample using Model (1). The results, reported in Column (2) of Table 6, show that the coefficient of *CS* remains significant and positive, confirming the robustness of our study's conclusions.

4.3.3. Consider firms' proprietary costs

Additional consideration is given to the determinants of firm information disclosure strategies to mitigate endogeneity concerns arising from sample self-selection bias. According to proprietary cost theory (Wagenhofer, 1990), firms incur proprietary costs when engaging in voluntary information disclosure. Firms experiencing higher proprietary costs are more likely to limit voluntary disclosures to prevent potential losses

Table 6
Robustness tests of heckman two-stage, PSM, considering firm's proprietary costs and DID.

Variables	Heckman	PSM	Low *PC*	High *PC*	DID
	(1) *TFP*	(2) *TFP*	(3) *TFP*	(4) *TFP*	(5) *TFP*
CS	0.148***	0.221***	0.255***	0.019	
	(3.57)	(4.99)	(4.93)	(0.41)	
IMR	−1.116**				
	(−2.38)				
Standard					0.096**
					(2.20)
Size	0.682***	0.562***	0.546***	0.564***	0.558***
	(13.14)	(30.99)	(27.01)	(27.01)	(35.27)
Lev	0.465**	0.866***	1.310***	0.279***	0.875***
	(2.16)	(8.79)	(12.13)	(2.63)	(10.00)
Soe	−0.222*	0.089**	0.045	0.086*	0.077**
	(−1.73)	(2.15)	(0.95)	(1.92)	(2.13)
Fixed	−1.334***	−1.150***	−2.495***	−0.639***	−1.253***
	(−11.75)	(−9.57)	(−8.11)	(−5.00)	(−11.84)
TobinQ	−0.003	0.008	−0.028**	0.011	−0.002
	(−0.27)	(0.59)	(−2.03)	(0.87)	(−0.21)
Growth	0.139***	0.150***	0.120***	0.088***	0.112***
	(6.00)	(4.41)	(4.56)	(2.97)	(5.12)
Age	−0.071	−0.067	−0.088*	0.022	−0.053
	(−1.42)	(−1.23)	(−1.76)	(0.35)	(−1.17)
Indep	0.012*	−0.002	0.001	0.002	−0.000
	(1.93)	(−0.53)	(0.06)	(0.43)	(−0.06)
Dual	−0.066**	−0.087***	−0.069**	−0.083**	−0.072***
	(−2.30)	(−2.87)	(−2.21)	(−2.42)	(−2.75)
Top1	0.002	0.003***	0.002*	0.004***	0.003***
	(1.16)	(3.02)	(1.82)	(3.25)	(3.53)
Board	0.051	0.001	0.085	0.079	0.037
	(0.48)	(0.01)	(0.71)	(0.86)	(0.43)
Roa	2.360***	2.589***	2.984***	2.238***	2.756***
	(11.26)	(11.18)	(11.19)	(8.76)	(13.81)
Constant	−5.910***	−4.935***	−3.913***	−5.635***	−4.664***
	(−7.46)	(−8.40)	(−6.92)	(−9.03)	(−10.66)
Year	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes
Observations	5800	2867	2997	2996	6005
Adj-R^2	0.748	0.752	0.768	0.785	0.753

from information leakage, leading to selective disclosure practices. Based on this theory, we examine whether stable customer relationships continue to enhance a firm's TFP within the subgroup characterized by lower proprietary costs (i.e., firms exhibiting a lower degree of sample self-selection) to validate the robustness of our conclusions. Following Ellis et al. (2012), we measure a firm's proprietary costs (*PC*) using the ratio of fixed and intangible assets to total assets and conduct subgroup regressions based on the median value of *PC*. As shown in Columns (3) and (4) of Table 6, the coefficient of *CS* remains significant and positive in the subsample with lower PC, thereby confirming the robustness of our findings.

4.3.4. Staggered DID method

To further validate the robustness of our conclusion, we leverage exogenous shocks that influence customer stability. From 2014 to 2016, the Chinese government launched pilot policy for "logistics standardization" in 32 cities, including Beijing, Shanghai and Guangzhou. Logistics standardization, by unifying the specifications of key logistics facilities and supporting standardized information systems, significantly enhances the efficiency of goods circulation and reduces logistics costs (Zhang et al., 2016). It not only helps firms maintain stable customer relationships but also supports the development of high-level modern supply chains (Wang and Tan, 2019). Therefore, the logistics standardization pilot policy serves as a reasonable exogenous policy shock affecting customer stability. We construct a staggered DID design in which *Standard* equals 1 for treated firms in pilot cities after policy implementation. As shown in Table 6 (Column (5)), the positive and significant coefficient on *Standard* (0.096, t value = 2.20) confirms that exogenous improvements in customer stability enhance a firm's TFP, with parallel trends validated in Appendix B. The policy impact appears to be predominantly short-term. In the initial phase, the policy effectively reduces firms' transportation costs, leading to rapid improvements in customer relationship stability and a significant positive effect on TFP. However, as firms gradually adapt to the new standards and absorb the policy benefits, a diminishing marginal effect is observed. Moreover, with lower transaction costs, firms may expand their business scope and customer base over time (Li and Lu, 2024), which in turn may dilute the long-term impact of the policy on customer stability.

4.3.5. Instrumental variable method

To address potential endogeneity concerns, particularly reverse causality, we implement a rigorous two-stage least squares estimation framework. Social culture, as an informal institution, significantly influences corporate behavioral decisions. As demonstrated by Boardman and Kato (2003), Confucianism exerts positive effects on firms' commercial practices and the continuity of supply chain cooperation. Regions with stronger Confucian influence exhibit greater trust and commitment among local firms toward their supply chain partners, thereby fostering more stable cooperative relationships. To capture this effect, we use the spatial density of historical Confucian temple relics within a 200-kilometer radius of each firm as a proxy for Confucian cultural penetration. This geographic-cultural proxy is then interacted with industry-level customer stability to construct our primary instrumental variable (*IV1*). The instrument meets the relevance requirement through the established relationship between Confucian culture and supply chain behavior, whereas its exogeneity derives from the historical fixedness of temple locations and the industry-average computation of customer stability, thereby effectively isolating any direct impact on firms' TFP.

We further implement a geographic identification strategy following Hu et al. (2024), using city-level terrain relief as a second instrumental variable. This topographic measure affects supply chain stability through transportation costs (relevance condition), whereas its exogenous nature as a time-invariant geographic characteristic satisfies the exclusion restriction. Given that terrain relief is a cross-sectional variable, we construct a new instrumental variable (*IV2*) by interacting terrain relief with spatial differences in customer stability relative to neighboring cities ($\Delta Stable$). The $\Delta Stable$ variable captures disparities in supply chain relationships between a firm and firms in neighboring cities, but it does not directly affect the firm's TFP. This supports the validity of the new instrument.

To address potential concerns regarding regional time trends, we incorporate province-by-year fixed effects into the model. The regression results for the two instrumental variables are reported in Columns (1) to (4) of Table 7. The first-stage F-statistics exceed the conventional threshold of 10. In the second stage, the Kleibergen–Paap rk LM statistics are significant at the 1 % level, and the Cragg–Donald Wald F-statistics exceed the Stock–Yogo critical value of 16.38 at the 10 % significance level, thereby rejecting the null hypotheses of under-identification and weak instruments. In addition, the coefficients of *IV1*, *IV2* and *CS* are all significant and positive. These results confirm that our main conclusions remain robust to concerns regarding endogeneity.

4.3.6. Change model

Furthermore, we employ the change model (Eq. (2)) to examine the relationship between changes in the independent and dependent variables, thereby mitigating concerns related to omitted variables. $\Delta TFP_{i,t}$ represents the change in TFP for firm i in year t, calculated as $TFP_{i,t}$-$TFP_{i,t-1}$, and $\Delta Stable_{i,t}$ denotes the same calculation. As demonstrated in Column (5) of Table 7, the coefficient of $\triangle CS$ is significant and positive, indicating that firms' TFP increases with its customer stability, thereby supporting the robustness of our conclusion.

$$\Delta TFP_{i,t} = \beta_0 + \beta_1 \Delta Stable_{i,t} + \beta_2 \sum \Delta Controls_{i,t} + Ind + Year + Area + \varepsilon_{i,t} \tag{2}$$

Table 7
Robustness tests of instrumental variable method and change model.

Variables	Confucian culture		Terrain relief		Variables	Change Model
	(1) *CS*	(2) *TFP*	(3) *CS*	(4) *TFP*		(5) △*TFP*
IV1	0.131***					
	(7.82)					
IV2			0.658***			
			(12.22)			
CS		0.794***		0.109*	△*CS*	0.032**
		(2.65)		(1.82)		(2.53)
Size	0.006	0.552***	0.008	0.558***	△*Size*	0.162***
	(1.02)	(33.53)	(1.59)	(35.79)		(4.41)
Lev	−0.084***	0.943***	−0.072***	0.885***	△*Lev*	−0.046
	(−2.88)	(10.43)	(−3.03)	(9.99)		(−0.67)
Soe	0.013	0.071*	0.014	0.081**	△*Soe*	−0.073
	(0.96)	(1.96)	(1.29)	(2.28)		(−1.44)
Fixed	0.218***	−1.409***	0.133***	−1.258***	△*Fixed*	−0.976***
	(6.48)	(−11.95)	(4.69)	(−12.05)		(−12.87)
TobinQ	−0.000	−0.003	−0.001	−0.002	△*TobinQ*	0.019***
	(−0.11)	(−0.21)	(−0.37)	(−0.20)		(3.90)
Growth	−0.104***	0.191***	−0.064***	0.118***	△*Growth*	0.283***
	(−10.90)	(4.80)	(−7.43)	(5.31)		(21.49)
Age	0.005	−0.056	0.020	−0.050	△*Age*	0.073
	(0.30)	(−1.27)	(1.46)	(−1.14)		(0.50)
Indep	−0.003***	0.002	−0.003***	−0.000	△*Indep*	−0.001
	(−3.06)	(0.67)	(−3.40)	(−0.02)		(−0.50)
Dual	0.001	−0.070***	0.006	−0.068***	△*Dual*	−0.001
	(0.08)	(−2.64)	(0.64)	(−2.66)		(−0.01)
Top1	0.000	0.003***	0.000	0.004***	△*Top1*	−0.000
	(0.35)	(3.54)	(0.84)	(3.71)		(−0.099)
Board	−0.047*	0.078	−0.071***	0.039	△*Board*	−0.010
	(−1.62)	(0.87)	(−2.94)	(0.45)		(−0.24)
Roa	0.474***	2.381***	0.306***	2.697***	△*Roa*	0.854***
	(5.68)	(10.02)	(4.41)	(13.90)		(7.43)
Constant	0.401**	−4.901***	0.540***	−4.537***	*Constant*	0.196***
	(2.43)	(−10.19)	(3.83)	(−9.80)		(2.76)
Year	No	No	No	No	Year	Yes
Industry	Yes	Yes	Yes	Yes	*Industry*	Yes
Province	No	No	No	No	*Province*	Yes
Province × Year	Yes	Yes	Yes	Yes		
Observations	6005	6005	6005	6005	*Observations*	4114
Adj-R^2	0.166	0.727	0.393	0.750	*Adj-R^2*	0.481
F	61.22***		149.49***			
Kleibergen-Paap rk LM		75.28***		278.75***		
Cragg-Donald Wald F		127.81[16.38]		2401.56[16.38]		

4.3.7. Other measurements of the TFP and CS

First, the OLS method serves as a basic approach for estimating TFP, although it is susceptible to endogeneity concerns. In contrast, the FE method can eliminate the influence of time-invariant individual characteristics, thereby yielding more robust results. Accordingly, *TFP* is replaced with *TFP_OLS* and *TFP_FE*, recalculated using the two aforementioned methods, respectively. Second, following Wang and Peng (2016), we substitute *CS* with the three newly constructed variables introduced earlier: *CS_dummy*, *LnCS* and *CS_3*. The results presented in Table 8 indicate that the coefficients of all independent variables remain significant and positive. These findings confirm that our conclusions hold even after modifying the measurements of TFP and customer stability.

Table 8
Other measurements of the TFP and CS.

Variables	(1)	(2)	(3)	(4)	(5)
	TFP_OLS	*TFP_FE*	*TFP*	*TFP*	*TFP*
CS	0.151***	0.161***			
	(3.89)	(4.03)			
CS_dummy			0.066***		
			(3.64)		
LnCS				0.091***	
				(4.31)	
CS_3					0.133*
					(1.89)
Size	0.765***	0.823***	0.559***	0.558***	0.564***
	(48.37)	(50.87)	(35.44)	(35.39)	(27.07)
Lev	0.896***	0.918***	0.886***	0.892***	0.673***
	(9.81)	(9.82)	(10.12)	(10.19)	(5.51)
Soe	0.074**	0.075**	0.076**	0.073**	0.046
	(2.00)	(2.00)	(2.12)	(2.04)	(1.01)
Fixed	−0.287***	−0.096	−1.288***	−1.294***	−1.167***
	(−2.69)	(−0.88)	(−12.08)	(−12.14)	(−8.69)
TobinQ	−0.013	−0.014	−0.002	−0.002	0.014
	(−1.07)	(−1.16)	(−0.20)	(−0.17)	(0.91)
Growth	0.114***	0.107***	0.121***	0.130***	0.132***
	(4.98)	(4.58)	(5.54)	(5.97)	(4.00)
Age	−0.054	−0.054	−0.055	−0.053	−0.013
	(−1.19)	(−1.17)	(−1.23)	(−1.20)	(−0.18)
Indep	−0.000	−0.000	0.000	0.000	0.001
	(−0.04)	(−0.04)	(0.09)	(0.15)	(0.36)
Dual	−0.083***	−0.084***	−0.075***	−0.074***	−0.065*
	(−3.13)	(−3.11)	(−2.83)	(−2.83)	(−1.71)
Top1	0.004***	0.004***	0.003***	0.003***	0.004***
	(3.86)	(3.92)	(3.51)	(3.51)	(2.81)
Board	0.084	0.102	0.042	0.045	0.119
	(0.92)	(1.09)	(0.49)	(0.52)	(1.12)
Roa	2.819***	2.872***	2.712***	2.687***	2.711***
	(14.27)	(14.31)	(13.74)	(13.69)	(10.16)
Constant	−7.167***	−7.909***	−4.479***	−4.786***	−5.744***
	(−13.63)	(−14.73)	(−8.86)	(−9.27)	(−8.62)
Year	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes
Observations	6005	6005	6004	6005	2915
Adj-R^2	0.832	0.845	0.754	0.754	0.755

4.3.8. Other robustness tests

First, we introduce additional control variables and firm fixed effects. To minimize the omission of factors that may influence a firm's TFP, we further control for three potential variables: firm cash flow level (*Cashflow*), measured as net cash flow from operating activities divided by total assets; the current ratio (*Liquid*), measured as current assets divided by current liabilities; and the compensation of the top three executives (*Pay*), measured as the natural logarithm of the total compensation of the top three executives. Moreover, based on Model (1), we incorporate firm fixed effects to control for time-invariant firm-level determinants. Given the potential multicollinearity among firm, industry and province fixed effects, we also estimate a model that includes only firm and year fixed effects.

Second, we account for lagged time effects. In line with Zhang et al. (2024), to address potential endogeneity arising from reverse causality, we regress Model (1) again by lagging the independent variable *CS* and the control variables by one and two periods (*CS_L1* and *CS_L2*, respectively) to test the robustness of our conclusions.

Third, we conduct subsample regressions. We first exclude samples from the COVID-19 pandemic period (2020–2022) to mitigate the potential impact of increased environmental uncertainty and supply chain disruption risks during this time. We then restrict the sample to firms in the manufacturing sector, which is particularly susceptible to supply chain stability issues due to its reliance on intermediate goods (Hertzel et al., 2008).

All regression results are presented in Table 9. Overall, the findings confirm the robustness and credibility of our conclusions.

Table 9
Other robustness tests.

	add CVs and firm FE			lagged time effect		subsample regressions	
Variables	(1) *TFP*	(2) *TFP*	(3) *TFP*	(4) *TFP*	(5) *TFP*	(6) *TFP*	(7) *TFP*
CS	0.139***	0.067***	0.070***			0.153***	0.186***
	(3.72)	(3.06)	(3.10)			(4.06)	(4.20)
CS_L1				0.142***			
				(3.14)			
CS_L2					0.111*		
					(1.93)		
Cashflow	0.479***						
	(3.33)						
Liquid	−0.024***						
	(−4.75)						
Pay	0.103***						
	(4.54)						
Size	0.529***	0.465***	0.475***	0.562***	0.537***	0.551***	0.575***
	(31.05)	(13.96)	(14.01)	(30.38)	(24.45)	(34.08)	(26.67)
Lev	0.735***	0.167*	0.117	0.868***	0.816***	0.880***	0.640***
	(7.32)	(1.95)	(1.24)	(7.97)	(6.36)	(10.03)	(6.44)
Soe	0.078**	−0.054	−0.050	0.053	0.013	0.092**	0.100**
	(2.18)	(−0.77)	(−0.73)	(1.28)	(0.25)	(2.54)	(2.31)
Fixed	−1.400***	−0.961***	−0.961***	−1.133***	−1.053***	−1.259***	−1.176***
	(−12.74)	(−9.63)	(−9.02)	(−9.20)	(−7.42)	(−11.78)	(−9.42)
TobinQ	−0.008	0.030***	0.034***	0.026*	0.033*	−0.002	−0.015
	(−0.70)	(3.68)	(3.75)	(1.82)	(1.93)	(−0.20)	(−1.33)
Growth	0.137***	0.209***	0.204***	0.157***	0.131***	0.132***	0.083***
	(6.25)	(13.10)	(12.75)	(5.12)	(3.14)	(5.47)	(2.81)
Age	−0.069	0.398**	0.392**	−0.052	−0.023	−0.063	0.060
	(−1.57)	(2.57)	(2.58)	(−0.97)	(−0.35)	(−1.41)	(1.19)
Indep	0.000	0.000	0.000	0.001	0.003	−0.001	−0.001
	(0.15)	(0.20)	(0.10)	(0.27)	(0.74)	(−0.21)	(−0.46)
Dual	−0.068***	−0.001	−0.013	−0.103***	−0.134***	−0.075***	−0.100***
	(−2.62)	(−0.04)	(−0.55)	(−3.25)	(−3.37)	(−2.84)	(−3.37)
Top1	0.004***	−0.000	−0.000	0.004***	0.005***	0.003***	0.004***
	(3.95)	(−0.10)	(−0.21)	(3.24)	(3.71)	(3.44)	(3.35)
Board	0.017	0.001	−0.001	0.035	0.048	0.054	0.062
	(0.20)	(0.02)	(−0.01)	(0.34)	(0.42)	(0.64)	(0.71)
Roa	2.163***	1.547***	1.430***	2.430***	2.100***	2.658***	2.529***
	(10.38)	(9.40)	(8.17)	(9.04)	(6.30)	(12.85)	(11.30)
Constant	−5.274***	−2.744***	−3.442***	−4.496***	−4.839***	−4.564***	−5.338***
	(−10.36)	(−2.88)	(−3.77)	(−8.66)	(−6.87)	(−9.91)	(−9.39)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	No	Yes	Yes	Yes	Yes
Province	Yes	Yes	No	Yes	Yes	Yes	Yes
Firm	No	Yes	Yes	No	No	No	No
Observations	6000	6005	6005	4113	2898	5269	3543
Adj-R^2	0.760	0.570	0.546	0.728	0.698	0.754	0.776

4.4. Mechanism

To assess how customer stability enhances a firm's TFP, we establish the following transmission mechanism test model to examine H2 and H3:

$$M_{i,t} = \beta_0 + \beta_1 CS_{i,t} + \beta_2 \sum Controls_{i,t} + Ind + Year + Area + \varepsilon_{i,t} \tag{3}$$

$$TFP_{i,t} = \gamma_0 + \gamma_1 CS_{i,t} + \gamma_2 M_{i,t} + \gamma_3 \sum Controls_{i,t} + Ind + Year + Area + \varepsilon_{i,t} \tag{4}$$

where $M_{i,t}$ in Eqs. (3) and (4) represents the mediating variable. Based on the theoretical analysis presented in Section 2.2, the mediating variables include AC and Rep. β_1 represents the impact of CS on the mediating variable, whereas the sign and magnitude of γ_2 are crucial to determine whether the proposed mechanism exists. If γ_2 is significant and positive, it indicates that CS considerably enhances the mediating variable, thereby improving TFP, and vice versa.

Table 10
Mechanism test.

Variables	Type I Agency Cost		Firm Reputation	
	(1) *AC*	(2) *TFP*	(3) *Rep*	(4) *TFP*
CS	−0.029***	0.080**	0.149*	0.106***
	(−4.18)	(2.41)	(1.85)	(2.99)
AC		−2.336***		
		(−15.69)		
Rep				0.124***
				(12.60)
Size	−0.010***	0.536***	2.223***	0.277***
	(−3.77)	(37.41)	(50.82)	(9.21)
Lev	−0.115***	0.621***	1.572***	0.799***
	(−7.65)	(8.05)	(10.87)	(9.28)
Soe	−0.013**	0.044	0.027	0.078**
	(−2.26)	(1.39)	(0.43)	(2.29)
Fixed	−0.038**	−1.384***	0.154	−1.310***
	(−2.43)	(−14.59)	(0.84)	(−12.32)
TobinQ	0.015***	0.033***	0.165***	−0.028**
	(5.83)	(3.19)	(7.19)	(−2.31)
Growth	−0.025***	0.067***	0.043	0.089***
	(−5.67)	(3.37)	(1.03)	(4.00)
Age	0.014	−0.022	0.167*	−0.065
	(1.65)	(−0.53)	(1.83)	(−1.57)
Indep	0.000	0.001	−0.003	0.000
	(0.70)	(0.46)	(−0.59)	(0.04)
Dual	0.010**	−0.050**	−0.164***	−0.062**
	(2.06)	(−2.08)	(−2.76)	(−2.52)
Top1	−0.000*	0.003***	0.001	0.003***
	(−1.87)	(3.20)	(0.67)	(3.36)
Board	0.020	0.091	0.081	0.026
	(1.35)	(1.17)	(0.51)	(0.32)
Roa	−0.404***	1.746***	13.162***	1.350***
	(−8.98)	(8.93)	(25.80)	(4.68)
Constant	0.375***	−3.864***	−47.317***	1.339*
	(4.81)	(−8.33)	(−42.46)	(1.84)
Year	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes
Observations	6005	6005	5366	5366
Adj-R^2	0.484	0.801	0.851	0.781
Sobel test	Z value = 5.635P value = 0.000		Z value = 2.431P value = 0.015	
Bootstrap test		[0.04963, 0.11070]		[0.06173, 0.14976]

As reported in Table 10, the coefficients of *CS* in Column (1) and *AC* in Column (2) are both significant and negative at the 1 % level, indicating that *CS* improves TFP by reducing Type I agency costs. Furthermore, the coefficients of *CS* in Column (3) and *Rep* in Column (4) are significant and positive at the 10 % and 1 % levels, respectively, suggesting that stable customer relationships enhance firm reputation and thus *TFP*. Additionally, as shown in Table 10, our mechanism testing results are consistent with those of Sobel and bootstrap tests. In summary, these findings support H2 and H3.

5. Additional analysis

5.1. Supply chain spillover effect

Customer stability arises from interactions between focal firms and their customers and may affect both parties involved in the contract. As key stakeholders in business activities, supply chain partners, namely customers and suppliers, produce a significant spillover effect across various dimensions, including TFP (Serpa and Krishnan, 2018), ESG performance (Tang et al., 2023) and digital transformation (Li et al., 2024). Accordingly, this section examines whether the impact of customer stability on a focal firm's TFP is transmitted downstream along the supply chain, potentially producing a positive spillover effect on the customer's TFP.

Table 11
Supply chain spillover effect and its mechanism.

Variables	Supply chain spillover effect			Supply chain finance mechanism			
	(1) *TFP_C*	(2) *TFP_C_OLS*	(3) *TFP_C_FE*	(4) *Payable*	(5) *Receivable*	(6) *Credit = 0*	(7) *Credit = 1*
CS	0.189***	0.158**	0.160**	0.043***	0.416*	0.228	0.181**
	(2.65)	(2.24)	(2.22)	(3.42)	(1.81)	(1.60)	(2.28)
C_Size	0.486***	0.685***	0.743***	−0.012***	0.112*	0.458***	0.492***
	(19.55)	(26.81)	(28.16)	(−2.87)	(1.80)	(10.17)	(17.47)
C_Lev	1.041***	0.987***	1.004***	0.191***	0.636	0.588*	1.162***
	(5.37)	(5.06)	(5.05)	(5.80)	(1.23)	(1.85)	(5.42)
C_Soe	0.020	0.023	0.026	0.012	−0.032	0.166	−0.031
	(0.35)	(0.42)	(0.45)	(1.10)	(−0.22)	(1.61)	(−0.54)
C_Fixed	−1.162***	−0.317	−0.145	−0.060	0.424	−1.463***	−1.039***
	(−5.42)	(−1.62)	(−0.73)	(−1.60)	(0.84)	(−5.79)	(−4.16)
C_TobinQ	0.018	0.011	0.012	0.002	0.123	−0.027	−0.015
	(0.53)	(0.33)	(0.34)	(0.32)	(1.52)	(−0.34)	(−0.51)
C_Growth	0.171*	0.186**	0.180*	0.014	0.508**	−0.042	0.236**
	(1.84)	(2.02)	(1.94)	(0.96)	(2.07)	(−0.27)	(2.33)
C_Age	0.319***	0.312***	0.312***	0.026	0.224	0.556***	0.210**
	(3.40)	(3.35)	(3.30)	(1.52)	(0.98)	(3.40)	(2.03)
C_Indep	−0.006	−0.005	−0.005	−0.001	0.014	0.003	−0.005
	(−1.44)	(−1.15)	(−1.13)	(−1.03)	(0.97)	(0.33)	(−0.98)
C_Dual	−0.013	−0.012	−0.007	−0.001	0.030	−0.015	−0.014
	(−0.24)	(−0.21)	(−0.12)	(−0.06)	(0.18)	(−0.15)	(−0.23)
C_Top1	0.000	−0.000	−0.000	0.001	0.012**	0.001	−0.000
	(0.08)	(−0.02)	(−0.06)	(1.61)	(2.28)	(0.29)	(−0.19)
C_Board	−0.136	−0.123	−0.113	−0.001	0.204	−0.148	−0.094
	(−0.97)	(−0.88)	(−0.79)	(−0.04)	(0.49)	(−0.68)	(−0.59)
C_Roa	2.979***	2.877***	2.897***	−0.003	−1.852	3.795***	2.774***
	(5.45)	(5.48)	(5.47)	(−0.03)	(−1.39)	(4.20)	(4.79)
Constant	−4.536***	−6.672***	−7.379***	0.373***	−4.256**	−2.232*	−3.873***
	(−6.05)	(−8.94)	(−9.68)	(2.65)	(−2.53)	(−1.87)	(−4.52)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	750	750	750	713	750	228	522
Adj-R^2	0.853	0.910	0.918	0.582	0.210	0.851	0.864

To empirically test the hypothesized spillover effect, we construct a one-to-one matched sample of "focal firm–customer–year" observations by manually pairing disclosed top five customer data from listed companies with basic firm information from the CSMAR database. We exclude non-listed and cross-listed companies from the customer sample, resulting in 750 observations. Subsequently, we establish Eq. (5):

$$TFP\_C_{i,t} = \delta_0 + \delta_1 CS_{i,t} + \delta_2 \sum C\_Controls_{i,t} + Ind + Year + Area + \varepsilon_{i,t} \tag{5}$$

where $TFP\_C_{i,t}$ represents the *TFP* of customer i in year t, which is calculated using LP, OLS and FE methods, respectively (*TFP_C*, *TFP_C_OLS* and TFP_C_FE, respectively). The independent variable remains a focal firm's *CS*, and the control variable is customers' characteristics.

The regression results are presented in Table 11. For *TFP* calculated using the LP, OLS and FE methods, the coefficients of *CS* are all significant and positive. This indicates that the impact of *CS* on a focal firm's *TFP* can generate a supply chain spillover effect and facilitate improvements in customers' *TFP*.

To further examine the potential transmission mechanism of this spillover effect, we consider the role of trade credit. Focal firms frequently extend trade credit to their customers as part of routine business operations, which can help alleviate customers' financing constraints. This, in turn, may enhance their operational efficiency (Zhang et al., 2024) and improve their TFP (Meza et al., 2019). Accordingly, we hypothesize that supply chain finance (specifically, trade credit) functions as a mediating mechanism in the observed spillover effect.

To test our hypothesis, we adopt a dual measurement approach from both the customer and focal firm perspectives. First, following the methodology described by Yu et al. (2024), we examine changes in the total trade credit (*Payable*) received by the customer. Second, to precisely identify bilateral trade credit flows at the focal firm–customer dyadic level, we construct a novel focal firm-specific trade credit measure (*Receivable*) that allocates trade credit proportionally based on each customer's share of the focal firm's total annual sales. This refined measurement strategy is theoretically grounded in established economic principles, which suggest that the trade credit extended to customers is positively correlated with their annual purchase volume within focal firm–customer relationships (Petersen and Rajan, 1997). As shown in Columns (4) and (5) of Table 11, regardless of whether the dependent variable is *Payable* or *Receivable*, the coefficient of *CS* is significant and positive. This indicates that *CS* increases the amount of trade credit provided by the focal firm to its customers.

$$Payable = \frac{(AccountsPayable + NotesPayable - Prepayments}{TotalAssets} \tag{6}$$

$$Receivable = \frac{CustomerPurchases}{Focalfirm\prime sAnnualSales} \times Focalfirm\prime sDownstreamTradeCredit \tag{7}$$

$$Focalfirm\prime sDownstreamTradeCredit = \frac{(AccountsReceivable + NotesReceivable)}{TotalAssets} \tag{8}$$

To rigorously examine the transmission mechanism of supply chain finance, we use the first-difference in trade credit (ΔTC) from focal firms to downstream customers to identify changes in trade credit availability. A negative ΔTC, which indicates tightened credit policies by focal firms, reduces the availability of trade credit to downstream customers, thereby constraining their *TFP* improvement and resulting in a statistically nonsignificant supply chain spillover effect. In contrast, a positive ΔTC, indicating an expansion of trade credit, significantly enhances the supply chain spillover effect. To implement this analysis, we construct a binary indicator variable (*Credit*) that equals 0 when $\Delta TC < 0$ and 1 when $\Delta TC \geq 0$. The cross-sectional results presented in Columns (5) and (6) of Table 11 demonstrate that the supply chain spillover effect is significant only for firms with increased trade credit ($Credit = 1$) and statistically nonsignificant for those with decreased trade credit ($Credit = 0$). These findings provide robust empirical evidence that the observed supply chain spillover effect is mechanistically driven by trade credit extended from focal firms to downstream customers.

5.2. Heterogeneity analysis

This study further explores the heterogeneity in the impact of *CS* on *TFP* from three perspectives: industry type, marketization degree and social trust level.

First, regardless of industry type, information asymmetry persists when firms establish contractual relationships with clients, necessitating reliance on observable signals to support decision-making. Affiliation with the high-tech industry serves as a critical signal and benchmark that facilitates organizational decisions. High-technology firms generally exhibit stronger technological innovation capabilities, which promote technical collaboration and information sharing across supply chain networks (Isaksson et al., 2016). Moreover, stable customer relationships not only provide reliable financing support to high-tech firms but also serve as important feedback channels for technological and product innovation (Tu et al., 2025), thereby more effectively contributing to improvements in TFP.

Second, regarding the degree of marketization, the process of marketization has significantly contributed to economic growth. Regions with higher levels of marketization typically exhibit more efficient market mechanisms and sound legal environments, facilitating open and transparent communication and collaboration among firms within the supply chain (Zou et al., 2024), thereby enhancing supply chain stability. Moreover, highly marketized regions experience more intense market competition and greater efficiency in resource allocation. In these contexts, stable customer relationships enable firms to more effectively access external resources and forecast market demand, thereby improving resource utilization and production efficiency.

Finally, with respect to social trust levels, trust in social relationships plays an essential role in economic development and social exchange. It is widely regarded as a core component of social capital and a foundation for economic transactions (Brockman et al., 2020). Customer stability essentially reflects a form of “relational transaction” that is facilitated by higher levels of regional social trust. In such regions, the information environment is more transparent, and firms within the supply chain are more willing to share information, collaborate on innovation and extend trade credit (Wu et al., 2014).

Based on the aforementioned analysis, we posit that the effect of *CS* on *TFP* is stronger for firms in high-tech industries and those located in regions with higher levels of marketization and social trust. Accordingly, we conduct a heterogeneity analysis based on industry classification and the levels of marketization and social trust in the regions where the firms are located. The regression results are shown in Table 12 and support our conjecture. The technological endowment of high-tech firms enables them to manage supply chain relationships more effectively. A favorable external market environment and institutional background are also essential for enhancing firms’ production efficiency. These findings further enrich the conclusions of our study and provide significant implications for management practices by both firms and governments.

6. Discussion and conclusions

6.1. Discussion

Amid deepening global economic integration and rapidly evolving business environments, firms face increasing competition and external uncertainties. In this context, maintaining stable customer relationships is an effective strategy for navigating external shifts. Based on a series of empirical tests, this study confirms that stable customer relationships, as a crucial relational resource, play an essential role in enhancing a firm’s TFP.

Specifically, this study uses a sample of A-share listed companies from 2008 to 2022 to systematically examine the impact of customer stability on firms’ TFP and its underlying mechanisms. First, we find that higher customer stability significantly enhances firms’ TFP, with substantial economic implications. Second, the mechanism analysis reveals that customer stability enhances TFP primarily by reducing Type I agency costs and enhancing firm reputation, thereby clarifying the link between the two. Furthermore, using matched supply chain data, we find that the customer stability of focal firms also positively influences their customers’ TFP, and evidence suggests that supply chain finance is the primary mechanism. Finally, we explore the heterogeneity of this impact across firms and find that it is more pronounced in high-tech industries and in regions with higher levels of marketization and social trust. The reason may be that high-tech firms possess stronger innovation capabilities and firms in regions with more developed market mechanisms and favorable transaction environments are better positioned to benefit from customer stability in improving TFP.

Table 12
Heterogeneity analysis.

Variables	Industry type		Marketization degree		Social trust degree	
	(1) Non-high-tech	(2) High-tech	(3) Low	(4) High	(5) Low	(6) High
CS	0.073	0.216***	0.065	0.204***	0.081	0.222***
	(1.35)	(4.37)	(1.33)	(4.04)	(1.61)	(4.44)
Size	0.554***	0.562***	0.542***	0.572***	0.551***	0.560***
	(25.39)	(23.85)	(24.04)	(28.87)	(22.89)	(27.56)
Lev	0.908***	0.864***	0.665***	1.080***	0.551***	1.177***
	(7.06)	(7.55)	(5.41)	(9.67)	(4.39)	(10.09)
Soe	0.056	0.098**	0.204***	−0.067	0.182***	−0.036
	(1.0771)	(1.97)	(4.10)	(−1.40)	(3.77)	(−0.68)
Fixed	−1.255***	−1.387***	−1.145***	−1.470***	−0.991***	−1.606***
	(−9.13)	(−9.21)	(−7.95)	(−10.72)	(−7.05)	(−10.78)
TobinQ	0.016	−0.018	0.004	−0.003	−0.004	0.003
	(0.73)	(−1.58)	(0.30)	(−0.19)	(−0.23)	(0.20)
Growth	0.184***	0.070**	0.106***	0.176***	0.136***	0.134***
	(6.06)	(2.19)	(3.70)	(5.53)	(4.43)	(4.55)
Age	−0.228***	0.074	−0.092	−0.020	−0.097	−0.019
	(−3.05)	(1.40)	(−1.40)	(−0.38)	(−1.25)	(−0.38)
Indep	0.002	−0.002	0.002	−0.001	0.004	−0.002
	(0.61)	(−0.61)	(0.46)	(−0.32)	(1.08)	(−0.69)
Dual	−0.066	−0.095***	−0.107***	−0.051	−0.060	−0.087***
	(−1.60)	(−2.96)	(−2.96)	(−1.54)	(−1.46)	(−2.72)
Top1	0.003**	0.003***	0.004***	0.002**	0.004***	0.003**
	(2.01)	(2.70)	(3.02)	(2.21)	(2.97)	(2.48)
Board	0.023	0.053	0.052	0.040	0.114	−0.008
	(0.16)	(0.56)	(0.51)	(0.29)	(1.107)	(−0.06)
Roa	2.80***	2.609***	2.57***	2.663***	2.681***	2.587***
	(10.32)	(10.06)	(10.20)	(10.04)	(10.02)	(9.72)
Constant	−3.509***	−4.303***	−4.196***	−5.148***	−4.900***	−4.096***
	(−4.84)	(−6.88)	(−6.15)	(−7.70)	(−7.97)	(−6.42)
Year	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2837	3168	2965	3040	2890	3115
Adj-R^2	0.760	0.757	0.774	0.754	0.781	0.751

6.2. Theoretical contributions

This study contributes to the literature in several ways. First, it extends research on the determinants of a firm's TFP. Previous studies examine the impact of supply chain finance (Gu et al., 2023) and blockchain innovation (Xu and Guan, 2023) on TFP from a supply chain perspective. However, despite the growing importance of ensuring the safety and stability of industrial and supply chains, no study has yet focused on the impact of supply chain customer relationships on TFP. This study addresses this gap by investigating the influence of customer stability on TFP, providing a novel perspective on improving firm production efficiency.

Second, this study examines the specific mechanisms through which customer stability affects firm TFP, focusing on both governance and reputation. This analysis not only helps to clarify the mechanisms underlying supply chain customer management effects but also contributes to research on the economic consequences of customer stability. Although numerous studies investigate the impact of stable supply chain relationships on firm performance (Baiman and Rajan, 2002; Fornell et al., 2006; Gosman and Kohlbeck, 2009) and the influence of customer stability on supply chain finance (Liu et al., 2022a) and trade credit (Zhang et al., 2024), dedicated studies examining its impact on production efficiency and the associated mechanisms are lacking. However, this study fills this gap by demonstrating that customer stability improves corporate governance and firm reputation, thereby influencing TFP by reducing Type I agency costs and bolstering firm reputation. These findings extend existing supply chain management theory.

Third, this study further contributes to the literature on the supply chain spillover effect. Serpa and Krishnan (2018) posit that firm productivity is influenced by three key factors: the firm's own characteristics, the productivity of its partners and the characteristics of its partners. Previous studies show that a firm's technological advancement and resource allocation efficiency (Hu et al., 2015), financial constraints (Hopenhayn, 2014) and the productivity of its supply chain partners (Serpa and Krishnan, 2018) all influence firm TFP. Building on this foundation, the present study shows that high customer stability not only enhances the TFP of focal firms but also increases the TFP of their customers through the extension of additional trade credit. This finding underscores that the productivity-enhancing effects of customer stability also produce a notable supply chain spillover effect, further contributing to the broader literature on the supply chain spillover effect.

6.3. Managerial implications

This study also offers valuable operational and practical insights for managers.

First, firms should prioritize customer relationship management as a strategic focus by enhancing their customer management systems to strengthen customer stability, thereby fostering a sustainable competitive advantage. In the current volatile environment, stable customers directly impact a firm's revenue generation and market demand forecasting, which in turn affects its profitability, risk-taking capacity and production efficiency (Gosman and Kohlbeck, 2009; Liu et al., 2022a). Accordingly, firms should invest in advanced customer relationship management systems and implement regular communication mechanisms to better monitor and analyze customer behaviors, needs and feedback. This allows for timely adjustments in products and services, thereby improving customer satisfaction and stability.

Second, firms should strengthen internal governance, reduce agency costs and enhance their social reputation. Agency costs and a poor social reputation are significant barriers to improving TFP (Chiang and Lin, 2007; Stuebs and Sun, 2010). Firms should establish effective corporate governance structures and risk management systems to ensure a clear allocation of responsibilities, improve internal transparency and enhance their risk-bearing capacity. Additionally, improving the quality of information disclosure can help reduce information asymmetry and build greater trust in the market.

Third, in the context of globalization and economic integration, firms along the supply chain should prioritize cooperation, increase investment in R&D and continuously strengthen their core competitiveness. This is consistent with the recommendations of Zhang et al. (2024). The findings of this study indicate that high-tech firms with strong technological innovation capabilities can better leverage customer stability to improve TFP. Therefore, while maintaining customer stability, focal firms should also actively pursue technological exchange and collaboration with firms in high-tech industries to promote product and service innovation and further improve TFP.

Fourth, it is important for relevant government management departments to further develop a market system characterized by orderly competition and comprehensive frameworks. They should focus on improving the business environment and advancing the marketization process. Additionally, the government can enhance social communication and information sharing to promote equity and justice within society, thereby fostering greater social trust. Such measures can further amplify the positive impact of customer stability on TFP.

6.4. Research limitations and future research

Although this study offers several theoretical contributions, it also has some limitations.

First, the measurement of customer stability is based solely on voluntarily disclosed "top five customers" information by firms, which may introduce data constraints. Future research could leverage big data or artificial intelligence to obtain more comprehensive customer information and derive a more accurate measure of customer stability.

Second, this study primarily focuses on the characteristics of stable customer relationships within the supply chain. Given the broad scope of supply and industrial chains, future studies could adopt a more comprehensive approach by examining additional aspects, such as customer concentration, supplier stability and concentration and the geographic distance and homogeneity between customers and suppliers.

Finally, the empirical evidence in this study is based on data from Chinese listed companies. While the focus on China's industrial and supply chain safety and stability provides valuable insights into enhancing firm TFP, it may impart our findings with distinctly Chinese characteristics, potentially limiting their generalizability to other economies. Future research could address this limitation by examining the impact of supply chain customer relationships on firm TFP in different economic environments.

Declaration of competing interest

We declare that there are no conflicts of interest regarding the research presented in this manuscript.

Acknowledgements

This work was supported by the Liaoning Province Social Science Fund (Grant No. L23BGL029).

Appendix A. Variable definitions

The Appendix presents the detailed definitions of variables.

Variable properties	Variables	Symbol	Description
Dependent variable	Total factor productivity	*TFP*	LP semi-parametric estimation method
Independent variable	Customer stability	*Stable*	The number of top five customers in the current year that also appeared in the previous year's top five, divided by five; the larger the value, the more stable the customers.
Mediating variables	Type I Agency Cost	*AC*	The ratio of management expenses to the total operating income
	Firm Reputation	*Rep*	Selection of 12 firm reputation evaluation indicators using factor analysis.
Control variables	Firm size	*Size*	The natural logarithm of the total assets at the year
	Asset-liability ratio	*Lev*	Total liabilities divided by total assets
	Firm profitability	*Roa*	Net profit divided by total assets
	Fixed asset ratio	*Fixed*	The ratio of fixed assets to total assets
	The percentage shareholding of the largest shareholder	*Top1*	The percentage shareholding of the largest shareholder
	Firm age	*Age*	The logarithm of numbers years since the establishment of the firm
	Nature of property rights	*Soe*	An indicator variable that equals 1 if the listed firm is state-owned, and 0 otherwise
	Firm relative value	*TobinQ*	The ratio of market value to the replacement cost of assets
	Operating income growth rate	*Growth*	The annual rate of growth in primary business income
	The percentage of independent directors	*Indep*	The proportion of independent board members
	Board size	*Board*	The natural logarithm of the total number of directors on the board
	The degree of separation between two rights	*Dual*	Difference between control rights and cash flow rights
	Year Fixed Effect	*Year*	Year dummy variable
	Industry Fixed Effect	*Ind*	Industry dummy variable
	Province Fixed Effect	*Province*	Province dummy variable

Appendix B. *Parallel trend test:*

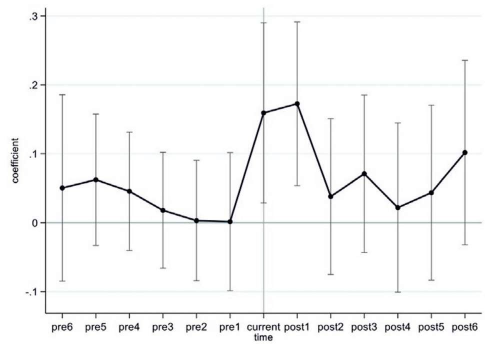


Fig. B1. Results of the parallel trend test.

References

Allen, F., Qian, J., Qian, M., 2005. Law, finance, and economic growth in China. J. Financ. Econ. 77 (1), 57–116.

Baier, S.L., Dwyer Jr, G.P., Tamura, R., 2006. How important are capital and total factor productivity for economic growth? Econ. Inq. 44 (1), 23–49.

Baiman, S., Rajan, M.V., 2002. Incentive issues in inter-firm relationships. Acc. Organ. Soc. 27 (3), 213–238.

Bauer, A.M., Henderson, D., Lynch, D.P., 2018. Supplier internal control quality and the duration of customer-supplier relationships. Account. Rev. 93 (3), 59–82.

Bellocchi, A., Sanchez Carrera, E.J., Travaglini, G., 2021. What drives TFP long-run dynamics in five large European economies? Economia Politica 38, 569–595.

Bernard, A.B., Moxnes, A., Saito, Y.U., 2019. Production networks, geography, and firm performance. J. Polit. Econ. 127 (2), 639–688.

Boardman, C.M., Kato, H.K., 2003. The Confucian roots of business Kyosei. J. Bus. Ethics 48, 317–333.

Bonacchi, M., Kolev, K., Lev, B., 2015. Customer franchise—A hidden, yet crucial, asset. Contemp. Account. Res. 32 (3), 1024–1049.

Brockman, P., El Ghoul, S., Guedhami, O., Zheng, Y., 2020. Does social trust affect international contracting? Evidence from foreign bond covenants. J. Int. Bus. Stud., 1–34

Chen, G., Tian, X.S., Yu, M., 2022. Redact to protect? Customers' incentive to protect information and suppliers' disclosure strategies. J. Account. Econ. 74 (1), 101490.

Cheng, Z., Zhu, Y., 2024. Does stock market liberalization increase company TFP? Evidence from the Shanghai-Shenzhen-Hong Kong Stock Connect program in China. Econ. Syst.

Chiang, M.H., Lin, J.H., 2007. The relationship between corporate governance and firm productivity: evidence from Taiwan's manufacturing firms. Corp. Gov. 15 (5), 768–779.

Ellis, J.A., Fee, C.E., Thomas, S.E., 2012. Proprietary costs and the disclosure of information about customers. J. Account. Res. 50 (3), 685–727.

Fornell, C., Mithas, S., Morgeson III, F.V., Krishnan, M.S., 2006. Customer satisfaction and stock prices: High returns, low risk. J. Mark. 70 (1), 3–14.

Freeman, R.E., Harrison, J.S., Wicks, A.C., 2007. Managing for stakeholders: survival, reputation, and success. Yale University Press.

Gosman, M.L., Kohlbeck, M.J., 2009. Effects of the existence and identity of major customers on supplier profitability: is Wal-Mart different? J. Manag. Account. Res. 21 (1), 179–201.

Gosman, M., Kelly, T., Olsson, P., Warfield, T., 2004. The profitability and pricing of major customers. Rev. Acc. Stud. 9, 117–139.

Gu, H., Yang, S., Xu, Z., Cheng, C., 2023. Supply chain finance, green innovation, and productivity: evidence from China. Pac. Basin Financ. J. 78 101981.

Guan, K.L., Zhang, R., 2019. Corporate reputation and earnings management: Efficient contract theory or rent-seeking theory. Account. Res. 375 (1), 59–64.

Guo, S., Zhang, Z., 2023. Green credit policy and total factor productivity: evidence from Chinese listed companies. Energy Econ. 128 107115.

Healy, P.M., Hutton, A.P., Palepu, K.G., 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. Contemp. Account. Res. 16 (3), 485–520.

Hertzel, M.G., Li, Z., Officer, M.S., Rodgers, K.J., 2008. Inter-firm linkages and the wealth effects of financial distress along the supply chain. J. Financ. Econ. 87 (2), 374–387.

Hofmann, E., Kotzab, H., 2010. A supply chain-oriented approach of working capital management. J. Bus. Logist. 31 (2), 305–330.
Hopenhayn, H.A., 2014. Firms, misallocation, and aggregate productivity: a review. Annu. Rev. Econ. 6 (1), 735–770.
Hu, C., Xu, Z., Yashiro, N., 2015. Agglomeration and productivity in China: firm level evidence. China Econ. Rev. 33, 50–66.
Hu, H., Bai, Z., Wang, A., 2024. Supply chain shareholding and high-quality development of enterprises: from the perspective of total factor productivity. China Ind. Econ. 09, 137–155.
Hui, K.W., Liang, C., Yeung, P.E., 2019. The effect of major customer concentration on firm profitability: competitive or collaborative? Rev. Acc. Stud. 24, 189–229.
Isaksson, O.H., Simeth, M., Seifert, R.W., 2016. Knowledge spillovers in the supply chain: evidence from the high-tech sectors. Res. Policy 45 (3), 699–706.
Kim, J.B., Song, B.Y., Zhang, Y., 2015. Earnings performance of major customers and bank loan contracting with suppliers. J. Bank. Financ. 59, 384–398.
Li, L.B., Lu, H.Y., 2024. Construction of the Unified National Market: how does logistics standardization promote domestic market integration? J. Financ. Econ. 50 (9), 19–33.
Lei, Q., Li, J., Zhong, Y., Huang, Y., 2023. Can differences in the background characteristics of the chairperson–CEO vertical dyad reduce management agency costs? —a perspective based on the internal configuration of the top management team. China J. Account. Res. 16 (1) 100293.
Levinsohn, J., Petrin, A., 2003. Estimating production functions using inputs to control for unobservables. Rev. Econ. Stud. 70 (2), 317–341.
Li, L., Bo, W., Qin, L., 2024. Spillover effect of digital transformation along the supply chain: from the perspective of suppliers' audit fees. China J. Account. Res. 17 (3) 100363.
Li, N., Yang, Z., 2011. Customer relationship and debt contracting. Work. Pap.
Liu, B., Ju, T., Chan, H.K., 2022a. The diverse impact of heterogeneous customer characteristics on supply chain finance: Empirical evidence from Chinese factoring. Int. J. Prod. Econ. 243 108321.
Liu, X., Liu, J., Wu, H., Hao, Y., 2022b. Do tax reductions stimulate firm productivity? A quasi-natural experiment from China. Econ. Syst. 46 (4) 101024.
Meza, F., Pratap, S., Urrutia, C., 2019. Credit, misallocation and productivity growth: a disaggregated analysis. Rev. Econ. Dyn. 34, 61–86.
Möller, K.K., Halinen, A., 1999. Business relationships and networks:: Managerial challenge of network era. Ind. Mark. Manag. 28 (5), 413–427.
Olley, G., Pakes, A., 1996. The dynamics of productivity in the telecommunications equipment industry. Econometrica 64, 1263–1297.
Panahifar, F., Byrne, P.J., Salam, M.A., Heavey, C., 2018. Supply chain collaboration and firm's performance: the critical role of information sharing and trust. J. Enterp. Inf. Manag. 31 (3), 358–379.
Patatoukas, P.N., 2012. Customer-base concentration: Implications for firm performance and capital markets: 2011 American accounting association competitive manuscript award winner. Account. Rev. 87 (2), 363–392.
Petersen, M.A., Rajan, R.G., 1997. Trade credit: theories and evidence. Rev. Financ. Stud. 10 (3), 661–691.
Saussier, S., 2000. Transaction costs and contractual incompleteness: the case of Électricité de France. J. Econ. Behav. Organ. 42 (2), 189–206.
Serpa, J.C., Krishnan, H., 2018. The impact of supply chains on firm-level productivity. Manag. Sci. 64 (2), 511–532.
Shen, B., Xu, X., Chan, H.L., Choi, T.M., 2021. Collaborative innovation in supply chain systems: Value creation and leadership structure. Int. J. Prod. Econ. 235, 108068.
Stuebs, M., Sun, L., 2010. Business reputation and labor efficiency, productivity, and cost. J. Bus. Ethics 96, 265–283.
Tang, J., Wang, X., Liu, Q., 2023. The spillover effect of customers' ESG to suppliers. Pac. Basin Financ. J. 78, 101947.
Tu, Y., Hu, L., Hua, X., Li, H., 2025. Supply chain stability and corporate green technology innovation. Int. Rev. Econ. Financ. 97, 103769.
Um, K.H., Kim, S.M., 2019. The effects of supply chain collaboration on performance and transaction cost advantage: the moderation and nonlinear effects of governance mechanisms. Int. J. Prod. Econ. 217, 97–111.
Wagenhofer, A., 1990. Voluntary disclosure with a strategic opponent. J. Account. Econ. 12 (4), 341–363.
Wang, C., Zhang, Y.J., 2024. Does environmental investment improve corporate productivity? Evidence from Chinese listed firms. Struct. Chang. Econ. Dyn. 70, 398–409.
Wang, K.L., Sun, T.T., Xu, R.Y., 2023. The impact of artificial intelligence on total factor productivity: empirical evidence from China's manufacturing enterprises. Econ. Chang. Restruct. 56 (2), 1113–1146.
Wang, X., Peng, X., 2016. Does stable relationship with customers improve the analyst forecasts about suppliers. J. Financ. Res. 5, 156–172.
Wang, X., Tan, J., 2019. Does logistics service standardization matter for corporate investment? China J. Account. Stud. 7 (4), 504–523.
Wernerfelt, B., 1984. A resource-based view of the firm. Strateg. Manag. J. 5 (2), 171–180.
Woo, D., Suresh, N.C., 2022. Voluntary agreements for sustainability, resource efficiency & firm performance under the supply chain cooperation policy in South Korea. Int. J. Prod. Econ. 252, 108563.
Wu, W., Firth, M., Rui, O.M., 2014. Trust and the provision of trade credit. J. Bank. Financ. 39, 146–159.
Xu, R., Guan, E., 2023. Can blockchain innovation promote total factor productivity? Evidence from Chinese-listed firms. Appl. Econ. 55 (6), 653–670.
Yang, Z., 2017. Customer concentration, relationship, and debt contracting. J. Appl. Acc. Res. 18 (2), 185–207.
Yu, Z., Zhao, X., Sun, L., 2024. Does supply chain voice influence firms' investment preferences? Financ. Res. Lett. 106055.

Zhang, J., Mo, H., Hu, Z., Zhang, T., 2024. The effect of stability and concentration of upstream and downstream relationships of focal firms on two-level trade credit. Int. J. Prod. Econ. 270, 109173.

Zhang, S., Luo, J., Huang, D.H., Xu, J., 2023. Market distortion, factor misallocation, and efficiency loss in manufacturing enterprises. J. Bus. Res. 154, 113290.

Zhang, Y., Liu, S., Liu, Y., Li, R., 2016. Smart box-enabled product–service system for cloud logistics. Int. J. Prod. Res. 54 (22), 6693–6706.

Zou, Y., Zhang, M., Zhang, M., 2024. The impact of company participation in supply chain alliances on the cost of equity capital: evidence from China. Int. Rev. Econ. Financ. 103387.