Distance makes a difference: Customer geographic proximity and suppliers' R&D investment intensity, do common institutional investors play a role?

Abstract

This study investigates whether and how customer geographic proximity affects innovation in supplier firms. By addressing customer concentration from the customer geographic distance aspect, we find that the geographic distance between a supplier and its main customers is negatively associated with supplier firm's R&D investment. Further analysis reveals that common institutional owners can effectively promote R&D investment due to their coordination function, eventually reaching a win-win situation within the supply chain. Interestingly, we find that customer geographic clustering, a dynamic setting of supplier-customer geographic proximity, positively moderates the relationship between customer geographic distance and supplier innovation investment.

Keywords: Customer Geographic Distance; R&D intensity; Customer Geographic Clustering, Common Institutional Investors

JEL classification: G14 G23 G34 M14

1. Introduction

There has been growing interest in supply chain partner relationships in academic research in recent years (Saboo et al., 2017; Chen et al., 2021). Existing studies demonstrate that customers play a crucial role in the supply chain, that affecting suppliers' decisions and behaviours significantly (Chu et al.,2019; Dai et al., 2021). Corporate innovation is crucial for firm development, offering firms competitive advantages and significant returns in the long run. However, there are limited evidence on the impact of customer geographic distance on supplier firms' innovation engagement, and the existing results remain controversial (Cao et al., 2023; Zheng et al., 2023). Consequently, it is worthwhile to explore how customer geographic distance affects innovation decisions in supplier firms.

In the past few years, global pandemic events like the COVID-19 have forced people to keep social distance and greatly boost the development of online communication tools like Zoom, DingDing, and VooV. Individuals increasingly rely more on those tools, raising questions about the necessity of geographic proximity. However, Peterson and Rajan (2002) document that the integral soft information for innovation is hard to communicate by digital tools or store electronically, it must rely on human interaction. Cao et al. (2022) argue that unforeseen black swan events create significant risks for corporate operations and management, geographical proximity reduces information asymmetry between customers and suppliers, thereby facilitating the trust and cooperation in supply chain. Therefore, firms consider adjusting the geographic distance with their supply chain partners to enjoy geographic proximity advantages. Ouyang et al. (2024) find that mutual funds and other business partners increase site visits to collect first-hand information due the fast development of high-speed railways in China[1]. Therefore, we expect that customer geographical distance affects suppliers' R&D investment intensity.

We propose two competing hypotheses given geographic distance paradox can be both beneficial and detrimental to suppliers (Chu et al., 2019; Wei and Sheng, 2023). Customer geographic proximity may positively promote R&D investment intensity in supplier firms. Geographical proximity is found to strengthen the knowledge spillover effects along the supply chain which promotes the local corporation's innovation and enhances knowledge transformation and acquisition (Hsu et al., 2022). Developing new markets far away from headquarters helps firms broaden the network and provide unique resource advantages when competing with their competitors (Dikova et al., 2016). However, it is also possible that customer geographic proximity may lead to several concerns that reduce the innovation engagement of supplier firms. Those concerns include knowledge homogeneity (Presutti et al., 2017), a lack of resources from different areas to strengthen firms' competitiveness (Dikova et al., 2016); and being easily affected by local incidents (Nath et al., 2010). Therefore, although geographic proximity offers advantages for face-to-face interactions, timely feedback, and soft information transmission, customer geographic distance may negatively impact R&D investment intensity in supplier firms.

As a fast-developing emerging economy, China provides us with multiple financial and political factors that are different from other developed economies, thereby motivating the

[1] Corporate site visits facilitate communication via in-person interactions (Huang and Fan, 2022).

proceed of this study. First, most of the existing studies on how customer geographic concentration affects supplier firms' innovation were conducted using the setting of developed economics (Azar et al., 2018; Kim and Zhu, 2018), few of them set targets for developing countries, while these markets are commonly less developed and lack efficient legal enforcement. Second, previous studies note that the top five customers take an average of more than 30% of annual sales in the Chinese A-share market (Chen et al., 2020; Zheng et al., 2023) which provides us with solid ground to investigate major customers' geographic proximity impact on suppliers' innovation intensities. In addition, China's setting provides us with multiple factors that may moderate the relationship between customer geographic proximity and corporate innovation intensity from the supply side, such as the emergence of state-owned common institutional investors.

Using a sample of A-share firms listed in the Shanghai Stock Exchange (SHSE) and Shenzhen Stock Exchange (SZSE) over the period from 2009 to 2021, a significantly negative relationship is found between customer geographic distance and suppliers' R&D investment intensity. That is, customer geographic distance reduces the innovation input of suppliers who are far away from their market destinations. Robustness and endogeneity tests are performed, and this result remains. Our results indicate that longer distances between suppliers and customers increase transaction costs, for instance, additional communication and operation expenditures, and require large amounts of input on relationship building (Buckley and Strange, 2011), further undercut the resources of firms' innovation. Distance also leads to communication barriers, making it harder to catch customers' timely needs and feedback, hindering the knowledge spillover effects between customers and suppliers (Chu et al., 2019; Wei and Sheng, 2023).

Interestingly, further analysis indicates that customer geographic clustering, a dynamic setting of supplier-customer geographic proximity, positively moderates the relationship between customer geographic distance and supplier R&D investment. This result support the argument of Bindroo et al. (2012) that customer geographic clustering promotes the connection between suppliers and customers, leading to higher customer involvement and real-time information flow between customers and suppliers. We also find that common institutional owners shape the relationship between customer geographic distance and supplier R&D investment. Our findings suggest that suppliers may attract common institutional investors to benefit from their coordination effect and knowledge-sharing ability, which moderate the negative relationship between customer geographic distance and supplier innovation investment intensity.

This study contributes to the literature in several ways. First, it enriches the studies on corporate innovation investment. Previous literature on the impact of customer concertation on suppliers' innovation mainly focus on sales concentration (Cao et al., 2023; Han et al., 2023; Zheng et al., 2023), this study indicates that customer geographic distance also matters. In addition, although literature has examined how customer concentration affects sustainable supply chain management and corporate risks (Krolikoeski and Yuan, 2017), evidence on the relationship between customer geographic distance and innovation is limited. Our paper, however, reveals how customers played in affecting upstream supplier innovation investment in emerging markets.

Second, this paper finds that customer geographic clustering positively moderates the negative impact of customer geographic distance on supplier R&D investment, providing further evidence of the dynamic impact of customer geographic proximity on innovation engagement in supplier firms. We argue that customer spatial clustering can affect supplier innovation intensity by enhancing the localized knowledge spillover effect. Our results, therefore, offer better understanding of the dynamic features of customer geographic distance.

Third, this paper adds evidence on the coordination and information sharing effects of common owners. We find that the negative impact customer geographic distance on supplier R&D investment is mitigated by common cross-holders. Common institutional investors relieve the stress from customers and provide suppliers with expertise guides in innovation engagement. They can work as a bridge to facilitate information exchange and help portfolio firms spend limited resources in the most necessary areas (He and Huang, 2017). Therefore, our evidence contributes to policymakers by highlighting a mechanism that promotes corporate innovation intensity through supply chains. For instance, our findings suggest that common owners can coordinate portfolio firms and enhance innovation input, ultimately achieving a win-win situation within the supply chain. Our paper is closely related to Chu et al. (2019) which document a positive relationship between supplier–customer geographic proximity and supplier innovation in the US. Compared to Chu et al. (2019), this paper further addresses customers' geographic clustering, and common institutional ownership matters in supply chain relationships and can positively moderate the negative effect of supplier-customer geographic distance. By enhancing the knowledge spillover effect, suppliers get verifiable and reliable soft information from spatial agglomerate major customers and collaborate closely with them under the guidance of common owners.

The rest of this paper proceeds as follows: Section 2 reviews the important related literature and proposes hypotheses. Section 3 describes the data set and sample construction. Section 4 presents the empirical results with various robustness tests, cross-sectional examinations, and further analysis. Section 5 concludes this study.

2. Literature Review and Hypothesis Development

Rich previous literature has investigated the influence of customer concentration on corporate decisions (Chu et al., 2019; Dai et al., 2021; Leung and Sun, 2021; Han et al., 2023). Supply chain concentration is an important factor that could affect corporate decisions, performance, and financial resilience along with supply chain partners (Dhaliwal et al., 2016; Chen et al., 2021). It is found that customer concentration endows suppliers with stronger resilience to ease shocks and higher inventory ratio, thereby avoiding excess inventory (Ak and Pataoukas, 2016; Hu et al., 2019; Chen et al., 2021). This study we particularity focuses on the relationship between customer geographic proximity and supplier firms' innovation intensity.

2.1 Customer Geographical Distance

This stream of literature argues that long-distance market destinations of customers are beneficial to corporate operations. Kim and Mathur (2008) note that regional-specific benefits could be granted with the network and economic scale development. As a key point of corporate expansion, developing new markets far away from headquarters helps firms broaden the network, provide unique resource advantages when competing with their competitors (Dikova et al., 2016), increase corporation resilience when target markets face unforeseen collapses (Nath et al., 2010; Liu et al., 2023). In addition, diversified markets far away might reduce the risk exposure when regional incidents occur, thereby strengthening corporate operational stability.

However, long-distance customer-supplier relationship requires high transaction costs and creates barriers to communication and idea transformation, which hinders the knowledge spillover effect to boost innovation (Chu et al., 2019). Long-distance customers from diverse communities may induce ethnic, religious, and ideological issues that lead to higher social distance, further increasing the cost of communication and difficulties in enjoying knowledge spillover. Based on transaction theory, Buckley and Strange (2011) argue that the expansion of the supply chain requires significant inputs on coordination, relationship, and network building.

Firms also need to consider idiosyncratic risk and their operation capability before they make customer expansion decisions. Liu et al. (2023) find that the benefits of market destination diversification will be undercut by additional inputs, such as spreading marketing and sales expenditure in diversified customer bases. A concentrated loyal customer base in few areas relaxes firms' pressure on extra information analysis, inputs into new customer preferences, and investigation of local legalization systems. Barbieri et al. (2020) investigate the effect of COVID-19 on customer proximity and conclude that supply chains with long-distance global market destinations across countries are much more vulnerable compared to those only concentrated in local markets. They also note that long-distance supply chain increases operational costs and risk exposure when facing uncertainties, leading to additional challenges for corporate management. In addition, the results of Barbieri et al. (2020) suggest that firms need to put recovering in the first place after suffering unexpected incidents which makes long-distance supply chain management unaffordable, especially in the post-COVID era. Customers in high proximity relationships prefer to participate in suppliers' corporate governance, restrain opportunistic behaviours, and continue to supervise enterprises to ensure the stability of supply chain relationships (Itzkowitz, 2015).

2.2 Customer Geographic Proximity and Corporate Innovation Intensity

Innovation is the whimsical process of bringing ideas into novel and use. For corporates, innovation includes a full series of activities from idea generation through input resources like research and development (R&D) expenses to commercializing concepts into outputs, such as products and services. Greater risks can be generated from non-produced results as the R&D input is manifested as sunk costs. Krolikoeski and Yuan (2017) explain that a reliable long-term partnership secures the needed resources for suppliers' innovation and reduces R&D expenditures by identifying their consumers' requirements given loyalty relationships often come from a concentrated customer base for those suppliers. Firms with sustainable supply chains can develop new products or innovate procedures to better capture market opportunities. In industry clusters, suppliers benefit from co-location with their agglomerated customers and generate externalities that boost innovation. The win-win business solutions are generated as a result of innovative strategies and the development of sustainable supply chains. Meanwhile, suppliers in a highly concentrated supply chain have stable sales channels with less market uncertainty which generates a stable external environment to maintain the smooth development of innovation sectors.

However, it is possible that suppliers with proximate customers no longer need to maintain strong R&D intensity and large-size innovation teams in the short term, as such they can invest more in operational sectors. Customer negotiation advantage motivates the knowledge transformation effect to suppliers which makes customers able to get competitive products with a lower price. Kikkawa et al. (2019) suggest that the direction of supply chain information transmission is driven by benefits and costs. They argue that when a firm faces falling input costs, other suppliers who sell to that business are also affected to lower their prices in the face of increased competition. Highly distance proximate customer groups increase the dependence on main customers and impair the corporation's ability to control the supply chain, also introducing more price and liquidity pressure on suppliers. This will damage corporate innovation ability in the long run and weaken suppliers' ability to independent innovation.

2.3 Hypothesis Development

The interfirm cooperation provides suppliers with innovative environments and knowledge authorizations when suppliers integrate their knowledge reservoir effectively (Wei and Sheng, 2023). Concentrated close-range markets could strengthen knowledge spillover effects which could promote the local corporation's innovation, and affect external knowledge transformation and acquisition (Hsu et al., 2022). Firms are more likely to have a direct and in-depth understanding of major customers' preferences when customer concentration is higher which makes their innovation activities more target-oriented (Delgado et al., 2020). Also, Barbieri et al. (2020) indicate that supply chains with long-distance market destinations are much more vulnerable compared to geographic proximate customers in local markets. In that case, the information transmission process will get blocked which makes focus on customers located nearby probably be a better choice for corporate innovation instead of broadening major customers far away.

On the contrary, regional-specific benefits could be granted with the broadening of distance networks and economic scale (Kim and Mathur, 2008). Multiple long-distance markets could bring unique resources and strengthen firms' competitiveness (Dikova et al., 2016), making their products and services hard to be copied by competitors. Also, these customers would not be affected by upstream firms' local incidents which increases suppliers' resilience when facing regional uncertainties (Nath et al., 2010). Furthermore, long-distance reduces the chance of knowledge homogenization and spatial lock-in, promoting suppliers to learn from their customers (Presutti et al., 2017). Based on previous findings, we expect that the impact of customer geographic proximity on supplier's corporate innovation intensity is under debate. Thus, the competing hypotheses are considered as follows:

H1a: Customer geographic proximity has a positive impact on corporate innovation intensity.

H1b: Customer geographic proximity has a negative impact on corporate innovation intensity.

The spatial agglomeration of customers can shape the relationship and characteristics along the supply chain (Bindroo et al., 2012) which represents that customer cluster can affect suppliers' innovation intensity geographically. Yang et al. (2024) discuss that when customers are more geographically proximate, their internal information transmission and aggregation are more likely to happen. Gertler (2003) notes that co-located people prefer to share a common code of communication and personal knowledge with higher trust in the information they exchange. In terms of geographic clusters, the cumulative spillover effects will make the information suppliers collected from their major customers more verifiable and reliable. Peterson and Rajan (2002) document that in-person interaction is an essential instrument for the transmission of soft information required for innovation. Meanwhile, geographically clustered locations increase the frequency and convenience of information sharing and exchange via face-to-face meetings and social relationships in both formal and informal channels. Therefore, concentrated customer geographic cluster has the communication advantage that coordinates suppliers' collection of integrated information, helps suppliers better understand the market through customers, and inspires innovation intensity. Hence, we expect that the customer geographic cluster plays a positive moderating role in the relationship between customer geographic proximity and corporate innovation intensity, and the second hypothesis is considered as follows:

H2: Customer geographic cluster positively moderates the relationship between customer geographic proximity and corporate innovation intensity.

Common institutional ownership, also known as common ownership, refers to the institutional investors holding more than 5% of outstanding shares in at least two firms in the same industry. Common institutional ownership was raised in the 1980s and was accompanied by increasing concentration among asset managers. Common ownership, as a key element of institutional ownership, plays an important monitoring role in China compared to other developed countries due to the concentrated state-owned shares, inexperienced individual investors, and the discrimination of minority ownership without government background (Liu et al., 2021).

Literature has discussed the coordination effect of common institutional ownership intensively. He and Huang (2017) point out there are at least two fundamental reasons why common owners improve the collaboration efficiency between firms in the same industry, that is incomplete contracting and information asymmetric. The coordination view believes that common ownership incentivizes the firms to cooperate more closely, facilitate forms of product market coordination, and improve operating profit margins and innovation productivity which benefits corporations through strategy channels (Azar et al.,2018; Lewellen and Lowry, 2021). The coordination effect of common ownership plays an important role in corporate mergers and acquisitions (M&As). It is found that common ownership can increase the merger likelihood between two firms, reduce the deal premium, and lead to better acquirer value (Brooks et al.,2018). Luong et al. (2017) note that institutional owners with long-term investment perspectives have the resources, abilities, and higher failure tolerance to guide innovation activities. Meanwhile, even minority M&As with small cash injection could raise the target firm's innovation due to the knowhow transferring from the acquirers to target firms (Boston and Spatareanu, 2018).

Common ownership may promote corporate innovation intensity due to efficient monitoring and better facilitating cross-firm knowledge transformation (Boston and Spatareanu, 2018). Firms with higher corporate governance levels are more likely to innovate while institutional owners are efficient monitors who increase firms' disclosure, improve portfolio firms' governance, and ultimately enhance the value of portfolio firms (Firth et al., 2016). To maximize portfolio benefits, common owners could enhance knowledge sharing and reduce information asymmetry to improve governance efficiency (He and Huang, 2017). In addition, Luong et al. (2017) provide evidence that corporate innovation is more likely to be guided by institutional investors who have more resources and higher failure tolerance. Similarly, common institutional investors have more expertise in specific industry sectors and may contribute to corporate innovation by directly promoting knowledge sharing with stronger power and indirectly creating a better innovation atmosphere. With the coordination effect from common owners, portfolio firms are expected to proceed with more innovation projects due to improved resource allocation and guidance from expert common investors.

As discussed, common owners promote knowledge sharing and information transparency (He and Huang, 2017). As institutional investors can better guide corporation innovation (Luong et al., 2017), the third hypothesis is as follows:

H3: Common ownership positively moderates the relationship between customer geographic proximity and corporate innovation intensity.

3. Data Description and Research Design

3.1 Sample Construction

The sample of this study contains A-share companies listed on the SHSE and SZSE, covering the period from 2009-2021. The sample period starts from 2009 because the supply chain information disclosure rate is quite low and the China Stock Market & Accounting Research

(CSMAR) database provides limited data on the supply chain for 2008 and before. In addition, the "Basic Specifications of Enterprise Internal Control" was first time put into practice in 2009 (CSRC, 2008), putting the disclosure of the source of revenues in the priority position, and providing us with more customer-side information to research. Most firms only have a limited number of key customers, and therefore, we focus on the top five disclosed customers following Chu et al. (2019) and Wei and Sheng (2023). Listed firms in financial industries and with missing variables are excluded from our sample set. The final panel data consists of 3938 firm-year observations from 2009 to 2021. All continuous variables are winsorized at the 1% and 99% level to reduce the effects from outliers.

3.2 Variable Construction

3.2.1 Customer geographic proximity and R&D intensity

Based on Dhaliwal et al. (2016) and Leung and Sun (2021), we construct measures to capture the supply chain proximity from the geographic distance aspect. To account for the major customers' geographic proximity across the supply chain, following Chu et al. (2019), and Wei and Sheng (2023), we collect the headquarters of firms and their major customers. We use the longitude and latitude coordinates with the consideration of the shape of Earth to calculate the distance between firm A and firm B in kilometers based on Kang and Kim (2008). With the concern of different numbers of disclosed major customers for sample firms, we follow Gaba and Meyer (2008), measuring *DIS1* as the weighted average of geographic distance between a firm and its customers' headquarters in a similar method. Specifically,

$$DIS1 = \sum_{j=1}^{j=5} Ln(Geographic\ Distance_{i,j,t}) \times\ Weight\ of\ Sales_{i,j,t} \qquad (2)$$

where *j* refers to the top three to five customers of firm *i* in year *t*. *Geographic Distance* refers to the distance collected from the CSMAR database. *Weight of Sales* indicates the percentage of total purchases of customer *j* divided by total sales of firm *i* in the year *t*.

We construct the corporate innovation intensity variables using the data obtained from the CSMAR database. Based on the annual R&D expense by the end of the year, we constructed two measures of innovation investment following the existing literature (Zhang and Kong, 2022). The first measure is the annual R&D expense divided by the firm's total assets at the end of the year (*RDasset*). We also construct the second measurement *RDsales*, measured by the ratio of R&D expense to total operating revenue by the end of the year.

3.2.2 Control Variables

Following Chu et al. (2019), we construct the control variables that might affect a firm's innovation intensity. The control variables include *Age* (the natural logarithm of firm age), *Size* (the natural logarithm of the firm total assets at the end of the year), *Lev* (the leverage of the firm which is measured by liability to total assets), *Cash* (cash holdings divided by total assets), *ROA* (net profit divided by total assets), *TQ* (Tobin's Q, measured by market value divided by the book value of total assets), *Tangibility* (asset tangibility which is calculated as property, plant, and equipment scaled by total assets), *SOE* (a dummy variable equals to 1 if the firm is state controlled, otherwise 0), *Largest* (the proportion of the largest shareholding), *BoardSize* (the natural logarithm of the number of board directors), and *BoardInd* (the ratio of independent directors to the total number of board directors). We also include common institutional ownership (*CIO1*) considering the high growth of the presence of common ownership in Chinese listed firms. *CIO1* is a dummy variable equal to 1 if institutional investors hold at least

5% of a firm's outstanding shares in at least two firms in the same industry for at least one quarter in a year, otherwise 0.

3.3 Summary Statistics

Table 1 provides summary statistics of the variables used in our study. Our sample firms invest an average of 1.9% of their total assets on R&D expenditure and 3.8% of their annual sales for innovation. These numbers are slightly different (higher) from typical numbers reported in previous literature due to the following reasons. First, we only focus on the supply chain partners' innovation input from the supplier side. Second, according to Banerjee et al. (2008) and Gao et al. (2024), relationship-specific investments can exist in supply chain relationships significantly which increases the possibility of R&D investments. The mean value of *DIS1* is 1.097 and the standard deviation is 1.143 with an average of 2.91 identifiable major customers disclosed by suppliers in our sample. All control variables are comparable to the previous studies. Specifically, common ownership has already existed in 9.1% of listed firms on average in China. Details of variable definition are available in Appendix A.

[Insert Table 1 here]

3.4 Model Specification

We use the following baseline model in Eqn. (3) to examine whether customer concentration has a significant impact on corporate innovation investments:

$$R\&D\ Intensity_{i,t} = \alpha + \beta Customer\ Geographic\ Proximity_{i,t} + \gamma Control_{i,t} + \varepsilon_{i,t} \quad (3)$$

where customer geographic proximity refers to geographic distance concentration for firm *i* in year *t*. All variables are defined as previously discussed, ε stands for the error term. We include firm and year fixed effects to control firm and time effects and firm's unobservable customer proximity heterogeneity. The correlation matrix is summarized in Table 2. The firm characteristics variables show a low level of correlation, representing that control variables are not highly correlated with each other.

[Insert Table 2 here]

4. Empirical Results

4.1 Baseline Results

Table 3 reports the impact of customer geographic proximity on supplier's innovation intensity. Columns 1 and 2 present the estimation of our baseline where *DIS1* is employed as the main independent variable when fixed supplier firms, respectively. We find that customer geographic distance is negatively associated with innovation intensity at the 5% significance level. Columns 4 and 5 present the baseline estimation results when controlling the industry-specific characteristics and the results remain negatively significant. Columns 3 and 6 present the baseline estimations based on lagged all independent variables for one period and the results remain negatively associated with customer geographic distance. Above results suggest that the corporation's R&D investments will decline significantly with the increase of customer geographic distance, respectively. According to the knowledge spillover theory, close-range markets promote R&D investment and external knowledge transformation (Hsu et al., 2022), while distant customers weaken the knowledge share effects, making suppliers less likely to benefit from localized knowledge spillover through the supply chain. Meanwhile, transaction cost theory suggests that the development of the distant market requires additional inputs on operation capability and information analysis to face the challenges in the strange place (Buckley and Strange, 2011). These resource inputs will undercut the diversification from the expansion and the resources planned to be allocated to innovation significantly (Sun and Govind, 2017; Liu et al., 2023). As a result, suppliers suffer from not only weakened financial resilience with extra costs but also have fewer disposable resources that could be assigned to innovation. Overall, the baseline results suggest that supplier innovation intensity is significantly and negatively affected by customer geographic distance.

[Insert Table 3 here]

4.2 Robustness Tests

This section reports the robust checks by using alternative dependent and independent variables. First, we alternatively measure the R&D intensity by employing *RDsales* following Zhang and Kong (2022). Column 1 in Table 4 shows that suppliers' innovation intensity remains significantly negatively affected by customer geographic distance which is consistent with our baseline results, suggesting that our baseline results are robust. Second, we employ *LnDis* calculated by the natural logarithm of the average customer's geographic distance plus one to measure customer geographic proximity alternatively following Wang et al. (2023). As shown in Column 2, *LnDis* remains significantly negatively associated with innovation expenditure, suggesting our baseline results still hold when using alternative independent variables. Consistent with the previous literature, geographically distinct customers causing hold-up problems can sabotage suppliers' operation strategies and lead to underinvestment in all departments (Krolikowski and Yuan, 2017). Suppliers also spread their limited resources in relationship maintenance with long-distance major customers, further reducing the cash flows that can be allocated to R&D. Overall, our robustness checks indicate that our baseline results are robust to alternative measures of both the dependent and independent variables.

[Insert Table 4 here]

4.3 Endogeneity

Endogeneity is a concern in this study due to reversed causality and unobservable firm specific factors that may affect innovation. For instance, manufacturing firms and information technology firms are naturally different in their operating behaviours, innovation strategies, and supply chain management. To address our industry-specific endogeneity concern, *DISmean* is employed following Dhaliwal et al. (2016). *DISmean* is defined as the average industry year of *DIS1* and serves as the instrumental variable (IV). Meanwhile, we employ the second instrumental variable *RDLS* which measures the relief degree of topography following Feng et al. (2007) as natural geographic factors heavily associated with population, firm distribution, and transport accessibility, further affecting in-person interaction and soft information transmission efficiency.

We then re-estimate Eqn. (3) and Table 5 presents the GMM estimation results of two-stage least squares (2SLS) analysis controlling for firm and year fixed effects. Columns 1 and 3 show the first stage results from the geographic distance aspect while Columns 2 and 4 show the second stage results. The first stage coefficients are all significant and positive while second stage coefficient result is significantly negative. The 2SLS tests indicate that our baseline results are robust when employing the GMM estimation and customer geographic distance negatively impacts suppliers' innovation intensity.

[Insert Table 5 here]

To alleviate the concern that firm characteristics or systematic differences may have impact on corporation R&D intensity which may lead to biased results. Hence, propensity-score matching (PSM) is used to address this concern by following Kong et al. (2020). We display a no replacement one-to-one nearest neighbour PSM method with the logit regression model and a caliper of 0.01. Our matched sample set contains 2669 observations for the geographic distance group as we determine independent variable by using *DIS2*. *DIS2* is a dummy variable equals to 1 if the distance is smaller than the median value of *DIS1*, otherwise 0. PSM results show that it is comparable for both treatment and control firm groups in the post-matched sample. We further re-estimate our baseline test by using Eqn. (3) with the matched data set and controlling firm and year fixed effects. The coefficients of *DIS1* in Table 6 Column 1 remain negative and significant which is consistent with our previous findings.

Meanwhile, we employ entropy balance (EB) following Ouyang et al. (2024) to adjust inequalities and balance the covariates of the sample set. Entropy balancing as a preprocessing scheme, directly adjusts the control variables and retains more information by using the covariate balance without counteracting bias reduction or losing observations. The balanced results based on three moments (mean, variance, and skewness) become similar instead of significantly different from each other suggesting that the balance match is effective. The re-estimated baseline coefficients shown in Column 2 are consistently significant and negative.

The above PSM and EB test results indicate that after accounting for the systematic differences, the adverse effect of customer geographic distance on suppliers' innovation intensity remains robust. Details of the PSM test and entropy balance results are presented in Appendix B.

[Insert Table 6 here]

Another concern is that the locations of different supply chain setups may lead to bias in our main findings as there are wide disparities in development between different regions of China. Hence, to address this concern we further control the provincial fixed effects based on our baseline estimations following Zhao and Wang (2024), and the results are presented in Table 7. As shown in Columns 1 and 2, customer geographic distance negatively affects supplier innovation intensity when controlling province factors at firm fixed level while Columns 3 and 4 identify the industry fixed effects. Table 7 results indicate that our baseline results are not biased by local specific factors and our findings hold.

[Insert Table 7 here]

4.4 Customer Geographic Clustering

When suppliers have multiple major customers, it is probable that several customers are located in the same region which rises the other dynamic, e.g., customer geographic clustering. Previous literature argues that interaction with distant major customers and assessment of their requirements become progressively more difficult (Ellis, 2007). Bindroo et al. (2012) suggest that major customers in the same spatial clusters can work as both competitors and collaborators, however, geographic clustering enhances the connection between suppliers and customers, leading to higher customer involvement and real-time information flow which may boost radical innovations. In this section, we further analyse the spatial agglomeration of customers along the supply chain by considering how customer geographic clustering shapes the relationship between customer geographic distance and supplier innovation.

Following Han et al. (2023), we employ two measures to proxy customer geographic clustering. Customer provincial concentration (*PC*) refers to customer special concentration and is defined as the number of major customers in the same province. Customer provincial dummy (*PD*) equals to 1 if at least 2 major customers are located in the same province, and 0 otherwise. As shown in Table 8 Columns 1 and 3, *PC* and *PD* are negatively related to suppliers' innovation intensity. Customers act as competitors due to sharing the same supplier in the same region, leading to the antipathy of information sharing to keep their competition superiorities. However, when considering the moderating effect of customer geographic clustering on the relationship between customer geographic distance and supplier innovation, the coefficients of the interaction terms *PC*DIS1* and *PD*DIS1* become positively significant. We also control for regional fixed effects, and our results hold as shown in Columns 2 and 4. The results suggest that customer spatial clustering exerts a positive impact on the relationship between customer geographic distance and supplier innovation. We argue that customer spatial clustering reduces suppliers' communication costs with major customers and enhances the likelihood of in-person interaction with major customers which is critical for soft information transmission. Consequently, customer geographic clustering positively moderates the relationship between customer geographic distance and supplier innovation.

[Insert Table 8 here]

4.5 Moderating Effect of Common Ownership

In this section, we explore the moderating factors of common ownership on the negative relationship between customer concentration and corporate innovation investments. Rich previous literature note that common ownership has the ability to enhance the coordination among holding firms, promote knowledge sharing, and reduce information asymmetry to maximize portfolio value (He and Huang, 2017; Azar et al., 2018). Considering the rapid development of common institutional ownership in the past decade, there are reasons to believe that common ownership may moderate the negative relationship between customer distance concentration and innovation investments.

To test the moderating effect of common ownership on geographic proximity, Column 1 in Table 9 shows that *CIO1*DIS2* is positively related to innovation inputs, suggesting that common ownership positively moderates the relationship between customer geographic distance and innovation intensity. In Column 2, a new dummy variable *City* measures whether suppliers and their largest customers in the same city is employed. The results of Column 2 show that *CIO1*City* is significantly positive, further confirming the positive moderating effect of common ownership. The results presented in Table 8 suggest that under the influence of common ownership, the negative impact of customer geographic distance on supplier innovation intensity can be mitigated.

[Insert Table 9 here]

Common owners could work as a bridge to facilitate information exchange and help portfolio firms identify their needs to spend limited resources in the most necessary areas (He and Huang, 2017). By providing cross-firm knowledge assistance for firms in high-competition environments with urgent knowledge needs, common owners can promote corporate innovation (Luong et al., 2017). Our evidence shows that common owners can coordinate portfolio firms and promote innovation investments to market competitiveness, eventually reaching a win-win situation that could maximize their benefits.

As the investigation deepens, common ownership is measured more rigorously by removing Hong Kong Securities Clearing Company Limited (HKSCC) from the sample and creating *CIO2*, given Jiang et al. (2022) argue that HKSCC cannot be seen as a common institutional investor. The HKSCC acts as a proxy corner for stockholders in the Hong Kong financial market, responsible for acting on behalf of equity holders and providing related services. It represents the foreign investors in China's mainland market and, should be removed from the common ownership calculation. Hence, we re-estimate the moderating effect test by excluding HKSCC as a common institutional investor and creating another dummy variable *CIO2,* which equals to 1 if the institutional investor excludes HKSCC holds over 5% of outstanding shares in at least two firms in the same industry, otherwise 0. The results are reported in Table 10. After redefining common institutional investors, the coefficients of interaction estimation remain statistically positively significant in both columns, suggesting that our results are less likely to be biased by the HKSCC. The negative relationship between customer geographic distance and corporation innovation intensity is positively moderated due to the coordination and knowledge-sharing of common owners, consistent with our previous findings.

[Insert Table 10 here]

4.6 Additional tests on customer R&D intensity and information sharing

Customer innovation intensity is another issue that needs to be addressed in this study. Hence, we employ two more variables *CI1* and *CI2* to further check the robustness of our baseline results. *CI1* is defined as the largest customer's innovation investment scaled by total sales, and *CI2* is defined as the largest customer's proportion of R&D personnel. As shown in Table 11 Panel A, the negative impact of customer geographic distance on suppliers' innovation intensity remains statistically significant. Innovative customers with large-size R&D departments have a higher requirement on their upstream suppliers' R&D investments with the potential to request competitive new products to maintain their own and entire supply chain market superiorities.

To further address the impact of state ownership on the supply chain information sharing process as SOEs take around 2/3 of industrial capital in China and the government background provides additional information exchange channels compared with privately owned firms (Liao et al., 2014). Hence, we performed a sub-sample test to identify the different impacts of geographic proximity on state-owned/private-owned enterprises. As presented in Panel B, the SOEs are not affected by the distance issues significantly. However, those non-SOEs are significantly negatively affected by the distance at both firm and industry level. Our results indicate that state ownership provides firms with unique information channels and helps SOEs ignore the negative impact of distance on their R&D intensity. Privately owned suppliers have to rely on vulnerable information transformation links which may hindered by geographically distant customers, hard to benefit from local knowledge spillover effects, and leads to recession of their innovation intensities.

Another consideration is that firm size also serves as an important factor that affects knowledge sharing and spillover effectiveness. Hence, we perform additional sub-sample tests based on the median firm size to further explore the impact. As shown in Panel C, relatively small suppliers suffer from a recession of R&D intensity significantly under the impact of customer distance while large-size suppliers do not. Petersen and Rajan (2002) point out that innovation requires soft information that heavily relies on in-person interaction instead of digital transformation, especially for small businesses. Relatively small-size suppliers tend to have less developed social networks as well as vulnerable information transformation links which makes them hard to take advantage of the local information spillover while long-distance increases their additional operation costs, further reducing the innovation intensity.

[Insert Table 11 here]

5. Conclusion

In response to a gradual increase in corporate innovation and supply chain research, this study focuses on analysing the influence of customer geographic distance on supplier innovation intensity and finds a negative association. Results show that longer distances from major customers result in lower suppliers' research and development inputs, due to the significant increase in relationship maintenance expenditures and weakened knowledge spillover effect. Our findings remain robust after addressing endogeneity concerns. In addition, common ownership positively moderates the negative association between customer geographic distance and corporate innovation intensity. A stricter measurement to exclude the HKSCC from the common ownership calculation has been performed and the results still hold. Common owners promote innovation intensity by using their coordination effect and promoting knowledge sharing to maximize their portfolio value. With the help and guidance from these industrial investment experts, firms not only get additional external knowledge about innovation and market trends but also avoid detours and twists in corporate decisions and investments. We also examine the customer spatial cluster effect find that proximate customer geographic cluster promotes supplier innovation due to enhanced likelihood of face-to-face interaction and better localized knowledge spillover effect. We also find that SOEs and large-size suppliers are not significantly negatively affected by customer geographic distance while privately owned and relatively small suppliers are due to lack of government connections and strong social networks.

This paper contributes to the growing literature on corporate innovation intensity and presents a new angle to investigate the moderating effects of common ownership in the finance literature. Different from previous literature, we detail identify the customer geographic proximity and its role in upstream suppliers' corporate decisions in emerging markets. We find that customer geographic distance still matters and significantly reduces the suppliers' innovation inputs while common ownership promotes R&D investments by information sharing and coordination ability to positively moderate the relationship between customer geographic distance and suppliers' innovation intensity, further contributing to the development of the geographic cluster of industry and integration of supply chain resources management.

References

Ak, B., Patatoukas, P. (2016). Customer-base concentration and inventory efficiencies: evidence from the manufacturing sector. *Production and Operations Management, 25(2),* 258-272.

Azar, J., Schmalz, M., Tecu, I. (2018). Anticompetitive Effects of Common Ownership. *The Journal of Finance, 73(4),* 1513-1565.

Bindroo, V., Mariadoss, B., Pillai, R. (2012). Customer clusters as sources of innovation-based competitive advantage. *Journal of Internation Marketing, 20 (3),* 17-33.

Brooks, C., Chen, Z., Zeng, Y. (2018). Institutional Cross-ownership and Corporate Strategy: The Case of mergers and acquisitions. *Journal of Corporate Finance, 48,* 187-216

Buckley, P., Strange, R. (2011). The governance of the multinational enterprise: insights internalization theory. *Journal of Management Studies, 48(2),* 460-470.

Campello, M., Gao, J. (2017). Customer concentration and loan contract terms. *Journal of Financial Economics, 123(1)*, 108-136.

Cao, F., Zhang, X., Yuan, R. (2022). Do geographically nearby major customers mitigate suppliers' stock price crash risk? *The British Accounting Review, 54(6),* 101118.

Chen, Y., Li, Q., Ng, J., Wang, C. (2021). Corporate financing of investment opportunities in a world of institutional cross-ownership. *Journal of Corporate Finance, 69,* 102041.

Chen, S., Lin, S., Xiao, J., Zhang, P. (2020). Do managers learn from stock prices in emerging markets? Evidence from China. *The European Journal of Finance, 28:4-5,* 377-396.

Cheng, X., Wang, H., Wang, X. (2022). Common institutional ownership and corporate social responsibility. *Journal of Banking & Finance, 136,* 106218.

Chu, Y., Tian, X., Wang, W. (2019). Corporate innovation along the supply chain. *Management Science, 65(6)*, 2455-2466.

Dai, R., Liang, H., Ng, L. (2021). Socially responsible corporate customers. *Journal of Financial Economics, 142,* 598-626.

Delgado, M., Mills, K. (2020). The supply chain economy: a new industry categorization for understanding innovation in services. *Research Policy, 49(8),* 104039.

Dhaliwal, D., Judd, J., Serfling, M., Shaikh, S. (2016). Customer concentration risk and the cost of equity capital. *Journal of Accounting and Economics, 61(1),* 23-48.

Dikova, D., Jaklic, A., Burger, A., Kuncic, A. (2016). What is beneficial for first-time SME-exporters from a transition economy: a diversified or a focused export-strategy? *Journal of World Business, 51(2),* 185-199.

Ellis, P. (2007). Distance, dependence and diversity of markets: effects on market orientation. *Journal of International Business Studies, 38(3),* 374-386.

Fan, D., Zhou, Y., Yeung, A., Lo, C., Tang, C. (2022). Impact of the U.S.-China trade war on the operating performance of U.S. firms: the role of outsourcing and supply base complexity. *Journal of Operations Management, 68(8),* 928-962.

Feng, Z., Tang, Y., Yang, Y., Zhang, D. (2007). Correlation between topographic relief and population distribution in China. *Acta Geographica Sinica, 62(10),* 1073-1082.

Firth. M., Gao. J., Shen. J., Zhang. Y. (2016). Institutional stock ownership and firms' cash dividend policies: Evidence from China. *Journal of Banking & Finance, 65,* 91-107.

Fu, Y., Qin, Z. (2021). Institutional cross-ownership and corporate philanthropy. *Finance Research Letter, 43,* 101996.

Gaba, V., Meyer, A. (2008). Crossing the organizational species barriers: How venture capital practices infiltrated the information technology sector. *Acadeour of Management Journal, 51(5)*, 976-998.
Gao, K., Shen, H., Gao, X., Chan, K. (2019). The power of sharing: Evidence from institutional investor cross-ownership and corporate innovation. *International Review of Economics and Finance, 63,* 284-296.
Gao, L., Han, J., Kim, J., Pan, Z. (2024). Overlapping institutional ownership along the supply chain and earnings management of supplier firms. *Journal of Corporate Finance, 84*, 102520.
Gertler, S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography 3 (1),* 75–99.
Han, Z., Yu, C., Li, Y., Shi, J., Liu, Y. (2023). Customer concentration and supplier financial performance: an inverted U-shaped relationship. *Industrial Marketing Management, 113,* 138-152.
He, J., Huang, J. (2017). Product Market Competition in a World of Cross-Ownership: Evidence from Institutional Blockholdings. *The Review of Financial Studies, 30(8),* 2674-2718.
Hsu, P., Hui, H., Lee, H., Tseng, K. (2022). Supply chain technology spillover, customer concentration, and product invention. *Journal of Economics & Management Strategy, 31,* 393-417.
Hu, H., Wang, L., Jiang, L., Yang, W. (2019). Strong ties versus weak ties in world-of-mouth marketing. *Business Research Quarterly, 22(4),* 245-256.
Huang, Y., Fan, Y. (2022). Risk along the supply chain: Geographic proximity and corporate risk taking. *Finance Research Letters, 50,* 103150.
Itzkowitz, J. (2015). Buyers as stakeholders: how relationships affect suppliers' financial constraints. *Journal of Corporate Finance, 31,* 54-66.
Jiang, P., Ma, Y., Shi, B. (2022). Common ownership and stock price crash risk: Evidence from China. *Australian Economic Papers, 61(4),* 876-912.
Kang, J., Kim, J. (2008). The geography of block acquisitions. *The Journal of Finance, 63(6),* 2817-2858.
Kikkawa, A., Magerman, G., Dhyne, E. (2019). Imperfect competition in firm-to-firm trade. *Available at SSRN 3389836,* National Bank of Belgium, Brussels.
Kim, D., Zhu, P. (2018). Supplier dependence and R&D intensity: The moderating role of network centrality and interconnectedness. *Journal of Operations Management, 64,* 7-18.
Kim, Y., Mathur, I. (2008). The impact of geographic diversification on firm performance. *International Review of Financial Analysis, 17(4),* 747-766.
Koch, A., Panayides, M., Thomas, S. (2021). Common ownership and competition in product markets. *Journal of Financial Economics, 139,* 109-137.
Kong, D., Wang, Y., Zhang, J. (2020). Efficiency wages as gift exchange: Evidence from corporate innovation in China. *Journal of Corporate Finance, 65,* 101725.
Krolikowski, M., Yuan, X. (2017). Friend or foe: Customer-supplier relationships and innovation. *Journal of Business Research,* 78, 53-68.
Lai, A., Chen, L., Wang, Q., Anderson, H. (2022). Natural disasters, trade credit, and firm performance. *Economic Modelling, 116,* 106029.
Leung, W., Sun, J. (2021). Policy uncertainty and customer concentration. *Production and Operations Management, 30(5)*, 1517-1542.
Lewellen, K., Lowry, M. (2021). Does common ownership really increase firm coordination? *Journal of Financial Economics, 141,* 322-344.

Liu, F., Liu, C., Wang, X., Park, K., Fang, M. (2023). Keep concentrated and carry on: redesigning supply chain concentration in the face of COVID-19. *International Journal of Logistics Research and Applications,* DOI: 10.1080/13675567.2023.2175803.

Liu, J., Wang, Z., Zhu, W. (2021). Does privatization reform alleviate ownership discrimination? Evidence from the Split-share structure reform in China. *Journal of Corporate Finance, 66,* 101848.

Luong. H., Moshirian. F., Nguyen. L., Tian. X., Zhang. B. (2017). How do foreign institutional investors enhance firm innovation? *Journal of Financial and Quantitative Analysis, 52,* 1449-1490.

Nath, P., Nachiappan, S., Ramanathan, R. (2010). The impact of marketing capability, operations capability and diversification strategy on performance: a resource-based view. *Industrial Marketing Management, 39(2),* 317-329.

Ouyang, C., Xiong, J., Liu, L., Yao, J. (2024). Geographic proximity and trade credit: Evidence from a quasi-natural experiment. *Journal of Corporate Finance, 84,* 102535.

Petersen, M. A., Rajan, R. G. (2002). Does distance still matter? The information revolution in small business lending. *The Journal of Finance, 57(6),* 2533-2570.

Presutti, M., Boari, C., Majocchi, A., Molina-Morales, X. (2017). Distance to customers, absorptive capacity and innovation in high-tech firms: The dark face of geographical proximity. *Journal of Small Business Management, 59(2)*, 343-361.

Saboo, A., Kumar, V., Anand. A. (2017). Assessing the impact of customer concentration on initial public offering and balance sheet-based outcomes. *Journal of Marketing, 81(5),* 42-61.

Wang, J. (2012). Do firm's relationship with principal customers/suppliers affect shareholders' income? *Journal of Corporate Finance, 18(4),* 860-878.

Wang, D., Liu, Z., Zhao, Z. (2016). Supply chain relationship and bank loan: based on supplier/customer concentration. *Accounting Research, 10,*42-49.

Wang, L., Wu, Y., Chen, Y., Dai, Y. (2023). Distance produces the fear of loss: customer geographic proximity and corporate cash holdings. *International Review of Financial Analysis, 87*, 102572.

Wang, Y., Liu, B., Chan, H., Zhang, T. (2023). Who pays buyer for not disclosing supplier lists? Unlocking the relationship between supply chain transparency and trade credit. *Journal of Business Research, 155(B),* 113404.

Wei, S., Sheng, S. (2023). Does geographic distance to customers improve or inhibit supplier innovation? A moderated inverted-U relationship. *Industrial Marketing Management, 108*, 134-148.

Yang, Y., Zhao, J., & Zou, G. (2024). Customer Geographic Clustering and Revenue Recognition: Evidence from China. *Journal of International Accounting Research, 23(2),* 201-223.

Zhang, D., Kong, Q. (2022). Credit policy, uncertainty, and firm R&D investment: A quasi-natural experiment based on the Green Credit Guidelines. *Pacific-Basin Finance Journal, 73,* 101751.

Zhao, T., Wang, H. (2024). The industry peer effect of enterprise ESG performance: the moderating effect of customer concentration. *International Review of Economics and Finance, 92,* 1499-1525.

Zheng, S., Zhang, Q., Zhang, P. (2023). Can customer concentration affect corporate ESG performance? *Finance Research Letters, 58(B),* 104432.

Table 1: Summary Statistics

This table presents the summary statistics of variables used in this study, covering 3938 firm-year observations from 2009-2021. All variables are defined in Section 3.2 and summarized in Appendix A.

Var	Obs	Mean	SD	Min	Median	Max
RDasset	3938	1.928	1.940	0.000	1.568	9.274
RDsales	3938	3.827	4.688	0.000	3.070	29.070
DIS1	3938	1.097	1.143	0.018	0.703	5.512
Age	3938	2.768	0.413	0.693	2.833	3.738
Size	3938	22.085	1.327	18.008	21.894	28.504
ROA	3938	0.032	0.064	-0.297	0.033	0.184
Lev	3938	0.431	0.216	0.050	0.425	0.922
TQ	3938	1.938	1.205	0.861	1.532	7.917
Tangibility	3938	0.222	0.160	0.006	0.190	0.690
CASH	3938	0.169	0.140	0.007	0.125	0.671
SOE	3938	0.203	0.402	0.000	0.000	1.000
Largest	3938	0.345	0.149	0.084	0.313	0.730
BoardSize	3938	2.265	0.171	1.792	2.302	2.773
BoardInd	3938	0.369	0.051	0.308	0.333	0.571
CIO1	3938	0.091	0.288	0.000	0.000	1.000
CIO2	3938	0.087	0.282	0.000	0.000	1.000
City	3938	0.212	0.407	0.000	0.000	1.000
PC	3938	1.340	0.815	1.000	1.000	5.000
PD	3938	0.228	0.423	0.000	0.000	1.000
CI1	679	0.029	0.031	0.000	0.027	0.249
CI2	679	0.068	0.105	0.000	0.000	0.660

Table 2: Correlation Matrix

This table presents the correlation matrix of variables used in this study. All variables are explained in Section 3.2 and summarized in Appendix A.

	RDasset	*DIS1*	*Age*	*Size*	*ROA*	*Lev*	*TQ*	*Tangibility*	*CASH*	*SOE*	*Largest*	*BoardSize*	*BoardInd*	*CIO1*
RDasset	1.0000													
DIS1	-0.0061	1.0000												
Age	-0.1065	0.0898	1.0000											
Size	-0.2851	-0.1100	0.2536	1.0000										
ROA	0.1601	-0.0822	-0.0770	-0.0483	1.0000									
Lev	-0.2170	0.0032	0.2390	0.5124	-0.4225	1.0000								
TQ	0.2714	0.1031	0.0268	-0.3949	0.0289	-0.1620	1.0000							
Tangibility	0.0588	0.0031	-0.1619	-0.0724	0.0749	0.0632	0.0024	1.0000						
CASH	-0.1434	-0.1262	0.1210	0.8114	0.1376	0.2652	-0.3668	0.0045	1.0000					
SOE	0.0121	-0.0713	0.0310	0.2364	-0.0575	0.2259	-0.0371	0.0839	0.1602	1.0000				
Largest	-0.0327	-0.0789	-0.1731	0.2046	0.0601	0.0132	-0.1020	0.0580	0.1948	0.0329	1.0000			
BoardSize	-0.0506	-0.1150	-0.0716	0.2570	-0.0206	0.1023	-0.1005	-0.1238	0.1616	0.2333	0.0884	1.0000		
BoardInd	0.0405	-0.0646	0.0168	-0.0201	-0.0140	0.0256	0.0222	0.0276	0.0353	-0.0571	-0.0381	-0.5586	1.0000	
CIO1	0.1170	-0.0591	0.0749	0.1406	-0.0153	0.0639	-0.0081	0.0326	0.1087	0.2345	0.0390	0.0967	-0.1062	1.0000

Table 3: Baseline Regression

This table presents the baseline regression results controlling for firm and year fixed effects to investigate the impact of customer geographic proximity on suppliers' R&D intensity. The regression model is as follows:

$RDasset_{i,t} = \alpha + \beta Customer\ Geographic\ Promixity_{i,t} + \gamma Control_{i,t} + \varepsilon_{i,t}$

All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	RDasset	*RDasset*	*$RDasset_{t+1}$*	*RDasset*	*RDasset*	*$RDasset_{t+1}$*
DIS1	-0.043*	-0.060**	-0.089**	-0.037*	-0.047**	-0.065**
	(-1.664)	(-2.364)	(-2.114)	(-1.707)	(-1.971)	(-2.231)
Age		-0.113	0.285		-0.259***	-0.065
		(-0.436)	(0.836)		(-3.196)	(-0.654)
Size		-0.368***	-0.047		-0.203***	-0.201***
		(-5.333)	(-0.483)		(-4.186)	(-3.403)
ROA		0.005	1.422***		2.268***	2.560***
		(0.016)	(3.328)		(5.005)	(4.381)
Lev		-0.316	-0.789***		-0.269	-0.487**
		(-1.589)	(-2.884)		(-1.576)	(-2.330)
TQ		0.040*	-0.051*		0.203***	0.120***
		(1.908)	(-1.798)		(8.243)	(3.630)
Tangibility		-0.077	0.175		0.285	0.133
		(-0.238)	(0.369)		(0.910)	(0.358)
CASH		-0.037	-0.017		0.142***	0.166***
		(-1.078)	(-0.380)		(3.732)	(3.616)
SOE		-0.103	0.551***		0.094	0.026
		(-1.031)	(2.631)		(1.326)	(0.336)
Largest		-0.350	-0.875*		0.380**	0.831***
		(-1.051)	(-1.945)		(1.965)	(3.602)
BoardSize		1.355***	2.091***		1.264***	1.823***
		(2.891)	(3.492)		(3.310)	(4.010)
BoardInd		0.198	1.583*		-0.404	-0.030
		(0.329)	(1.906)		(-0.712)	(-0.040)
CIO1		0.276***	0.007		-0.018	-0.109
		(3.254)	(0.068)		(-0.189)	(-1.004)
constant	1.916***	10.095***	0.394	1.939***	2.528***	1.045
	(6.029)	(6.219)	(0.172)	(5.869)	(2.924)	(0.961)
N	3938	3938	2147	3938	3938	2147
Firm FE	Yes	Yes	Yes	No	No	No
Industry FE	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.815	0.832	0.849	0.346	0.351	0.417

Table 4: Alternative Measure of R&D Intensity and Geographic Proximity

This table presents the results using alternative measure of R&D intensity. *RDsales* is defined as the proportion of R&D intensity on total annual sales at the end of the year. *LnDis* is defined as the natural logarithm of customer's average distance plus one following Wang et al. (2023). All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) *RDsales*	(2) *RDasset*
DIS1	-0.172*	
	(-1.778)	
LnDis		-0.031*
		(-1.895)
Age	-3.302***	0.882***
	(-3.345)	(7.154)
Size	0.241	-0.408***
	(0.917)	(-6.140)
ROA	-11.319***	-0.227
	(-9.077)	(-0.702)
Lev	-3.999***	-0.459**
	(-5.296)	(-2.422)
TQ	-0.235***	-0.015
	(-2.956)	(-0.782)
Tangibility	2.806**	-0.027
	(2.293)	(-0.083)
CASH	-0.546***	-0.041
	(-4.188)	(-1.223)
SOE	-0.084	-0.262***
	(-0.221)	(-3.177)
Largest	-0.773	-0.420
	(-0.610)	(-1.299)
BoardSize	-0.651	1.392***
	(-0.365)	(3.043)
BoardInd	2.541	0.658
	(1.107)	(1.079)
CIO1	0.692**	0.279***
	(2.149)	(3.496)
constant	18.883***	8.354***
	(3.060)	(6.155)
N	3938	3938
Firm FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.848	0.834

Table 5: Endogeneity: IV- GMM estimation

This table presents the results of two-stage least squares tests. *DISmean* and *RDLS* are employed as instrumental variables. *DISmean* is defined as the average industry-year of *DIS1*, *RDLS* is defined as the relief degree of topographic following (Feng et al., 2007). The first stage results are presented in Columns (1) and (3), the second stage results are presented in Columns (2) and (4). All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	DIS1	*RDasset*	*DIS1*	*RDasset*
DISmean	0.527***			
	(9.183)			
RDLS			1.671*	
			(1.608)	
DIS1		-0.003**		-1.901***
		(-2.441)		(-3.747)
Age	-0.265	-0.005***	-0.251	0.165
	(-1.376)	(-3.211)	(-1.252)	(1.192)
Size	-0.146***	-0.002***	-0.221***	-0.202*
	(-3.340)	(-3.740)	(-4.151)	(-1.797)
ROA	0.108	0.016**	0.170	-2.931**
	(0.414)	(2.232)	(0.671)	(-2.199)
Lev	0.138	-0.003	0.095	-1.476***
	(0.933)	(-0.907)	(0.621)	(-4.036)
TQ	0.030*	0.003***	0.023	0.350***
	(1.761)	(4.657)	(1.443)	(5.262)
Tangibility	-0.034	0.004**	-0.796***	-0.076
	(-0.454)	(2.308)	(-3.197)	(-0.123)
CASH	0.316**	0.019***	0.026	0.096
	(1.988)	(4.007)	(0.985)	(1.071)
SOE	-0.032	-0.000	-0.020	-0.412***
	(-0.426)	(-0.213)	(-0.255)	(-3.275)
Largest	-0.235	-0.003	-0.174	-0.450
	(-0.939)	(-0.899)	(-0.673)	(-1.357)
BoardSize	-0.170	0.002	-0.262	-1.379*
	(-0.979)	(0.661)	(-0.722)	(-1.936)
BoardInd	-0.135	0.002	-0.242	-1.673
	(-0.285)	(0.252)	(-0.520)	(-1.497)
CIO1	-0.066	0.002	-0.122*	0.405***
	(-1.046)	(0.988)	(-1.704)	(2.873)
constant	4.332***	0.062***	-1.142	8.342***
	(3.907)	(5.001)	(-1.608)	(5.034)
N	3938	3938	3938	3938
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Adj. R^2	0.732	0.147	0.729	0.080
F-Value	113.27		19.62	

Table 6: PSM Matched Sample and Entropy Balance Analysis

This table presents the results of propensity-score matching (PSM) and entropy balance (EB) tests. Column 1 shows the regression results using matched sample set containing 2,669 observations for the matched group. Column 2 shows the regression results using entropy balanced sample. All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) PSM *RDasset*	(2) EB *RDasset*
DIS1	-0.045*	-0.051*
	(-1.72)	(-1.951)
Age	-0.002	-0.187
	(-0.01)	(-0.477)
Size	-0.331***	-0.398***
	(-4.58)	(-3.138)
ROA	0.580	-0.244
	(1.47)	(-0.509)
Lev	-0.319	-0.146
	(-1.49)	(-0.479)
TQ	0.020	0.011
	(0.95)	(0.238)
Tangibility	0.136	-0.092
	(0.41)	(-0.163)
CASH	-0.051	-0.004
	(-1.43)	(-0.095)
SOE	-0.035	-0.179
	(-0.35)	(-1.178)
Largest	-0.693	-0.228
	(-1.99)	(-0.380)
BoardSize	0.908*	0.420
	(1.90)	(0.643)
BoardInd	0.051	-0.976
	(-0.08)	(-1.413)
CIO1	0.290***	0.131
	(3.27)	(1.555)
constant	8.744***	11.542***
	(5.42)	(3.654)
N	2669	3938
Firm FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.823	0.868

Table 7: Province fixed effect

This table presents the baseline regression results controlling for province characteristics. All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	*(3)*
	RDasset	*RDasset*
DIS1	-0.061**	-0.045*
	(-2.396)	(-1.833)
Age	-0.111	-0.261***
	(-0.427)	(-3.219)
Size	-0.369***	-0.202***
	(-5.342)	(-4.168)
ROA	0.002	2.275***
	(0.007)	(5.017)
Lev	-0.314	-0.268
	(-1.582)	(-1.574)
TQ	0.040*	0.203***
	(1.903)	(8.233)
Tangibility	-0.076	0.273
	(-0.236)	(0.868)
CASH	-0.037	0.141***
	(-1.082)	(3.719)
SOE	-0.102	0.092
	(-1.013)	(1.290)
Largest	-0.346	0.378*
	(-1.037)	(1.954)
BoardSize	1.361***	1.264***
	(2.902)	(3.310)
BoardInd	0.202	-0.401
	(0.334)	(-0.706)
CIO1	0.275***	-0.017
	(3.248)	(-0.181)
constant	10.099***	2.535***
	(6.221)	(2.934)
N	3938	3938
Firm FE	Yes	No
Industry FE	No	Yes
Province FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.821	0.351

Table 8: Customer geographical clustering

This table presents the regression results of the moderating effect of customer geographical clustering. *PC* refers to customer provincial concentration and is defined as the number of major customers in the same province. *PD* equals to 1 if at least 2 major customers are located in the same province, otherwise 0. All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	RDasset	*RDasset*	*RDasset*	*RDasset*
DIS1	-0.144***	-0.114**	-0.092***	-0.067**
	(-3.249)	(-2.567)	(-3.325)	(-2.415)
PC	-0.176***	-0.171***		
	(-3.515)	(-3.414)		
*DIS1*PC*	0.069***	0.065**		
	(2.709)	(2.534)		
PD			-0.217**	-0.217**
			(-2.385)	(-2.396)
*DIS1*PD*			0.168***	0.163***
			(3.276)	(3.185)
Age	-0.263***	-0.263***	-0.261***	-0.263***
	(-3.245)	(-3.205)	(-3.226)	(-3.207)
Size	-0.206***	-0.193***	-0.203***	-0.190***
	(-4.245)	(-3.929)	(-4.191)	(-3.864)
ROA	2.316***	1.977***	2.327***	1.987***
	(5.115)	(4.376)	(5.135)	(4.397)
Lev	-0.261	-0.171	-0.263	-0.173
	(-1.532)	(-0.999)	(-1.546)	(-1.012)
TQ	0.202***	0.201***	0.201***	0.200***
	(8.206)	(8.161)	(8.174)	(8.115)
Tangibility	0.334	0.293	0.337	0.305
	(1.065)	(0.926)	(1.074)	(0.965)
CASH	0.141***	0.118***	0.140***	0.115***
	(3.722)	(3.076)	(3.677)	(3.005)
SOE	0.096	0.181**	0.093	0.179**
	(1.345)	(2.465)	(1.307)	(2.435)
Largest	0.383**	0.436**	0.392**	0.443**
	(1.985)	(2.248)	(2.030)	(2.283)
BoardSize	1.261***	1.364***	1.283***	1.392***
	(3.306)	(3.558)	(3.363)	(3.631)
BoardInd	-0.445	-0.333	-0.381	-0.280
	(-0.784)	(-0.583)	(-0.672)	(-0.491)
CIO1	-0.009	0.008	-0.014	0.003
	(-0.093)	(0.080)	(-0.147)	(0.030)
constant	2.814***	2.814***	2.548***	2.558***
	(3.243)	(3.217)	(2.949)	(2.937)
N	3938	3938	3938	3938
Industry FE	Yes	Yes	Yes	Yes
Province FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Adj. R^2	0.353	0.368	0.352	0.368

Table 9: Moderating Effect: Common Ownership: Information Sharing & Coordination

This table presents the regression results of the moderating effect of common ownership. The *DIS2* is a dummy variable equals to 1 if the distance is smaller than the median value of *DIS1*. A dummy variable *City* is employed to further measure the geographic proximity and equals to 1 if the supplier and the largest customer are in the same city, otherwise 0. All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)
	RDasset	*RDasset*
CIO1	0.084	0.202**
	(0.747)	(2.327)
DIS2	0.140***	
	(2.971)	
*CIO1*DIS2*	0.306**	
	(2.330)	
City		-0.084
		(-0.349)
*CIO1*City*		0.315***
		(-4.941)
Age	-0.094	-0.201
	(-0.392)	(-0.663)
Size	-0.329***	-0.437**
	(-5.159)	(-2.377)
ROA	-0.162	0.035*
	(-0.537)	(1.827)
Lev	-0.442**	0.041
	(-2.410)	(0.138)
TQ	0.036*	-0.041
	(1.844)	(-1.281)
Tangibility	0.035	-0.090
	(0.119)	(-0.967)
CASH	-0.038	-0.200
	(-1.184)	(-0.649)
SOE	-0.087	1.307***
	(-0.937)	(3.011)
Largest	-0.238	0.162
	(-0.772)	(0.291)
BoardSize	1.374***	-0.084
	(3.176)	(-0.349)
BoardInd	0.193	-0.315***
	(0.346)	(-4.941)
constant	8.897***	8.786***
	(5.965)	(5.883)
N	3938	3938
Firm FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.828	0.827

Table 10: Alternative Measure of Common Ownership

This table presents the moderating effects of common ownership using an alternative measurement. Following Jiang et al. (2022), *CIO2* is constructed by removing Hong Kong Securities Clearing Company Limited (HKSCC) from the sample set. All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)
	RDasset	*RDasset*
CIO2	0.076	0.204**
	(0.662)	(2.341)
DIS2	0.141***	
	(2.998)	
*CIO2*DIS2*	0.327**	
	(2.439)	
City		-0.047
		(-0.872)
*CIO2*City*		0.289**
		(2.194)
Age	-0.093	-0.083
	(-0.387)	(-0.347)
Size	-0.330***	-0.316***
	(-5.175)	(-4.949)
ROA	-0.163	-0.202
	(-0.537)	(-0.665)
Lev	-0.438**	-0.435**
	(-2.387)	(-2.368)
TQ	0.035*	0.035*
	(1.830)	(1.825)
Tangibility	0.034	0.040
	(0.116)	(0.134)
CASH	-0.037	-0.040
	(-1.170)	(-1.273)
SOE	-0.089	-0.090
	(-0.961)	(-0.973)
Largest	-0.233	-0.198
	(-0.757)	(-0.640)
BoardSize	1.377***	1.308***
	(3.181)	(3.014)
BoardInd	0.198	0.163
	(0.355)	(0.292)
constant	8.900***	8.790***
	(5.968)	(5.886)
N	3938	3938
Firm FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.832	0.828

Table 11: Further Analysis

Panel A presents the regression results controlling for customers' innovation intensity. CI1 is defined as the largest customer's innovation investment scaled by total sales, CI2 is defined as the largest customer's proportion of R&D personnel. Panel B presents subsample regression results based on ownership type. SOE is a dummy variable equals to 1 if the firm is state-controlled, and otherwise 0. Panel C presents the sub-sample results based on median firm size. All variables are defined in Section 3.2 and summarized in Appendix A. *, **, and *** indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Panel (A) Customer innovation intensity

	(1)	(2)
	RDasset	*RDasset*
DIS1	-0.101*	-0.102*
	(-3.324)	(-3.239)
CI1	0.003	
	(0.141)	
CI2		0.011*
		(1.749)
ROA	3.525***	3.503***
	(3.295)	(3.285)
Lev	0.127	0.108
	(0.321)	(0.274)
TQ	0.192***	0.185***
	(3.107)	(2.994)
Tangibility	0.433	0.302
	(0.466)	(0.326)
CASH	0.202**	0.196**
	(2.247)	(2.182)
SOE	0.403**	0.413**
	(2.465)	(2.532)
Largest	1.220***	1.253***
	(2.620)	(2.696)
BoardSize	2.281**	2.192**
	(2.467)	(2.375)
BoardInd	0.774	0.685
	(0.604)	(0.536)
CIO1	0.181	0.158
	(0.877)	(0.768)
constant	3.978*	4.024*
	(1.855)	(1.881)
N	679	679
Industry FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.402	0.405

Panel (B) Ownership type

	SOE	non-SOE
	(1)	(2)
	RDasset	*RDasset*
DIS1	-0.013	-0.068**
	(-0.293)	(-2.185)
Age	-1.120**	-0.080
	(-2.055)	(-0.264)
Size	0.229*	-0.569***
	(1.804)	(-6.702)
ROA	0.762	-0.375
	(1.216)	(-0.981)
Lev	-0.675*	-0.244
	(-1.698)	(-1.051)
TQ	0.037	0.053**
	(0.833)	(2.214)
Tangibility	0.811	-0.252
	(1.280)	(-0.668)
CASH	-0.157**	-0.010
	(-2.443)	(-0.230)
Largest	1.030*	-0.920**
	(1.665)	(-2.239)
BoardSize	0.006	1.849***
	(0.007)	(3.254)
BoardInd	-1.273	1.001
	(-1.229)	(1.331)
CIO1	0.298**	0.326***
	(2.229)	(3.015)
constant	2.295	13.431***
	(0.801)	(6.741)
N	799	3139
Firm FE	Yes	Yes
Year FE	Yes	Yes
Adj. R^2	0.866	0.815

Panel (C) Firm size

	Large	Small
	(1)	(2)
	RDasset	*RDasset*
DIS1	0.020	-0.119***
	(0.747)	(-2.883)
Age	0.490*	-0.395
	(1.663)	(-0.816)
Size	-0.040	-0.614***
	(-0.488)	(-4.530)
ROA	0.560	-0.979**
	(1.287)	(-2.149)
Lev	-0.822***	-0.211
	(-3.278)	(-0.727)
TQ	0.020	0.032
	(0.535)	(1.075)
Tangibility	0.987***	-1.208**
	(2.594)	(-2.194)
CASH	-0.069*	-0.031
	(-1.778)	(-0.554)
SOE	0.025	-0.589**
	(0.287)	(-2.551)
Largest	-0.306	-0.028
	(-0.916)	(-0.045)
BoardSize	0.395	2.275***
	(0.814)	(2.875)
BoardInd	0.097	-0.014
	(0.163)	(-0.013)
CIO1	0.219***	0.392**
	(2.752)	(2.247)
constant	1.550	16.158***
	(0.789)	(5.353)
N	1969	1969
Firm FE	Yes	Yes
Year FE	Yes	Yes
Adj. R2	0.882	0.775

Appendix A. Variable Definition

Variable Name	Description
Dis1	The weighted average geographic distance between a firm and its customers' headquarters locations, following Wei and Sheng (2023).
LnDis	The natural logarithm of the average customer's geographic distance plus one following Wang et al. (2023).
City	Dummy variable equals to 1 if the firm and its largest customer are registered in the same city, otherwise 0.
RDasset	R&D intensity divided by total assets at the end of the year.
RDsales	R&D intensity divided by total operating revenue at the end of the year.
Age	Age of the firm, calculated by the natural logarithm firm age.
Size	Natural logarithm of total assets.
Lev	Leverage, measured by liability to total assets.
Cash	Cash holding divided by total assets.
ROA	Return on assets, measured by net profit to total assets
TQ	Tobin's Q, calculated by market value divided by the book value of total assets.
Tangibility	Asset tangibility, calculated by PPE scaled by total assets.
SOE	Dummy variable equals to 1 if the firm is state controlled, otherwise 0.
CIO1	Dummy variable, equals to 1 if institutional investors who hold at least 5% of a firm's outstanding shares in at least 2 firms in the same industry over one quarter, otherwise 0.
CIO2	Dummy variable, equals to 1 if institutional investors (not include HKSCC) who hold at least 5% of a firm's outstanding shares in at least 2 firms in the same industry over one quarter, otherwise 0.
Largest	Ratio of total shares held by the largest shareholder.
BoardSize	Natural logarithm of the number of board directors.
BoardInd	Ratio of number of independent board directors to the total number of directors.
PC	Provincial concentration, calculated as thenumber of major customers located in the same province with suppliers.
PD	Provincial dummy, dummy variable equals to 1 if at least 2 major customers located in the same province with suppliers, otherwise 0.
CI1	The largest customer's innovation investment scaled by total sales.
CI2	The largest customer's proportion of R&D personnel

Appendix B. Balanced tests after PSM

This table shows detail results of PSM and entropy balancing. We display a no replacement one-to-one nearest neighbour propensity scores method with the logit regression model and a caliper of 0.01. Our matched sample set contains 2,669 observations for the matched group while the unmatched sample contains 3938 observations. The balanced results are presented in Panel A. Pre-matched and post-matched results are presented in Panel B. Comparison of means before and after entropy balancing are presented in Panel C. All variables are defined in Section 3.2 and summarized in Appendix A.

Panel (A)

Variable	Sample	Treated	Control	%bias	|bias|	t-stat	p>|t|
Age	U	2.741	2.795	-13.1		-4.12	0.000
	M	2.760	2.770	-2.4	81.5	-0.69	0.491
Size	U	22.132	22.030	7.9		2.47	0.014
	M	22.096	22.083	1	86.9	0.29	0.772
ROA	U	0.038	0.025	20.3		6.38	0.000
	M	0.036	0.036	-0.6	97.3	-0.19	0.846
LEV	U	0.438	0.424	6.6		2.06	0.040
	M	0.428	0.431	-1.1	82.9	-0.32	0.749
TQ	U	1.881	1.995	-9.5		-2.97	0.003
	M	1.902	1.897	0.4	95.6	0.12	0.903
Tangibility	U	0.234	0.209	15.6		4.88	0.000
	M	0.218	0.218	0.3	98.3	0.07	0.941
CASH	U	0.174	0.163	7.4		2.33	0.020
	M	0.172	0.170	1	86.2	0.29	0.774
SOE	U	0.211	0.195	4.2		1.31	0.191
	M	0.208	0.194	3.5	16.6	0.98	0.328
Largest	U	0.352	0.337	10.4		3.26	0.001
	M	0.346	0.344	1.5	85.1	0.43	0.664
BoardSize	U	2.272	2.258	8.5		2.66	0.008
	M	2.266	2.264	1.2	85.6	0.34	0.731
BoardInd	U	0.370	0.369	0.9		0.29	0.769
	M	0.368	0.369	-1.3	-38.4	-0.37	0.713
CIO1	U	0.075	0.108	-11.3		-3.54	0.000
	M	0.089	0.082	2.4	78.5	0.70	0.483

Panel (B)

	DIS2	
	pre-match	post-match
Age	-0.216**	0.100
	(-2.50)	(0.66)
Size	-0.014	0.062
	(-0.38)	(1.93)
ROA	4.156***	-0.473
	(6.63)	(-1.96)
Lev	0.994***	-0.009
	(4.71)	(-0.09)
TQ	-0.062**	-0.011
	(-1.99)	(-0.84)
Tangibility	0.852***	-0.134
	(4.11)	(-2.74)
CASH	0.490*	-0.096
	(1.78)	(-0.91)
SOE	0.083	0.003
	(0.98)	(0.07)
Largest	0.274	-0.198
	(1.17)	(-1.14)
BoardSize	0.448*	0.005
	(1.92)	(0.04)
BoardInd	1.218	-0.094
	(1.61)	(-0.27)
CIO1	-0.394**	0.071
	(-3.37)	(1.57)
constant	-1.350***	-0.989
	(-1.42)	(-1.08)
N	3938	2669
Firm FE	Yes	Yes
Year FE	Yes	Yes
Pseudo. R2	0.024	0.575

Panel (C): Entropy balancing: Comparison of before and after entropy balancing

	Before entropy balancing							After entropy balancing						
	Treated n=921			Control n=3017				Treated n=921			Control n=3017			
	Mean	Variance	Skewness	Mean	Variance	Skewness	Std. Diff.	Mean	Variance	Skewness	Mean	Variance	Skewness	Std. Diff.
Dis1	1.347	1.419	1.661	1.020	1.247	1.897	0.176***	1.347	1.419	1.408	1.347	1.419	1.408	0.000
Age	2.884	0.140	1.408	2.747	0.167	-0.836	0.413***	2.844	0.140	-1.145	2.844	0.140	-1.145	0.000
Size	22.290	1.743	-1.145	22.020	1.644	0.698	0.112*	22.290	1.743	0.558	22.290	1.744	0.558	0.000
ROA	0.028	0.004	0.558	0.033	0.004	-1.942	-0.318	0.028	0.004	-2.144	0.028	0.004	-2.144	0.000
LEV	0.455	0.047	-2.144	0.424	0.046	0.226	-0.022	0.455	0.047	0.141	0.455	0.047	0.142	0.000
TQ	1.917	1.503	0.142	1.944	1.438	2.578	0.025	1.917	1.503	2.411	1.917	1.503	2.411	0.000
Tangibility	0.917	0.010	2.411	0.927	0.007	-2.502	-0.744*	0.917	0.010	-2.182	0.917	0.010	-2.183	0.000
CASH	20.260	1.757	-2.182	20.110	1.629	0.274	-0.000	20.260	1.757	0.143	20.260	1.757	0.143	0.000
SOE	0.240	0.183	0.143	0.192	0.155	1.567	0.183*	0.240	0.183	1.218	0.240	0.183	1.217	0.000
Largest	0.347	0.224	0.443	0.344	0.022	0.556	-0.042	0.346	0.022	0.443	0.346	0.022	0.443	0.000
BoardSize	0.937	0.007	-0.181	0.935	0.007	-0.188	-0.284	0.937	0.007	-0.181	0.937	0.007	-0.181	0.000
BoardInd	0.368	0.002	1.450	0.370	0.003	1.589	-0.490	0.368	0.002	1.450	0.368	0.002	1.450	0.000
CIO1	0.122	0.107	2.316	0.082	0.075	3.042	0.224*	0.122	0.107	2.316	0.122	0.107	2.314	0.000