Learning from Customers: Corporate Innovation along the Supply Chain

Yongqiang Chu, Xuan Tian, and Wenyu Wang*

Current Version: October, 2014

Abstract

This paper studies the effect of supplier-customer relationship on supplier innovation through a knowledge spillover channel. We use the geographical distance between a supplier and its major customers to capture knowledge spillovers along the supply chain. To establish causality, we explore plausibly exogenous variation in distance caused by customer headquarters relocations. In a difference-in-differences framework, we show that knowledge spillovers from customers appear to have a positive, causal effect on supplier innovation. The effect is stronger when the customers are more innovative themselves and are within closer technology proximity with the suppliers. Finally, we show that innovation attributable to knowledge spillovers from customers positively contributes to a firm's product market performance. Our paper sheds new light on the real effect of knowledge spillovers along the supply chain - its enhancement on firm innovation.

Key Words: Innovation, Knowledge Spillovers, Supplier-Customer Relationship, Supply Chain

*Chu is with Moore School of Business, University of South Carolina; Tian and Wang are with Kelley School of Business, Indiana University. We thank Gustavo Manso, Sophie Shive (discussant) and participants in the 2014 State of Indiana Conference for helpful comments.

1 Introduction

A growing literature has examined various effects of supplier-customer relationship on corporate decisions.[1] While most existing studies highlight the importance of the interactions between suppliers and customers along the supply chain in corporate finance, these studies focus on how supplier-customer relationships affect financial decisions. The existing literature has largely ignored an important impact of supplier-customer relationships: its real effect on corporate investment decisions. In this paper, we focus on a special type of corporate investment – technological innovation, which is critical for a firm's long-term competitive advantages and sustainable growth (Porter, 1992), and explore a key underlying channel – knowledge spillovers – through which supplier-customer relationship affects innovation.

Supplier-customer relationship could affect corporate innovation through knowledge spillovers in several different ways. First, a close relationship between a supplier and its major customers enable the supplier to learn the specific needs of its customers and hence stimulates more research and development (R&D) spending on the part of the supplier to satisfy its customer needs, which ultimately leads to technological innovation of the supplier (Han, Kim, and Srivastava, 1998; Lukas and Ferrell, 2000). Manso (2011) develops a model on mechanisms that motivate exploration (such as technological innovation) versus exploitation (such as routine tasks) and shows that timely feedback on the performance to the agent is critical for motivating innovation. A close relationship between a supplier and its customers also allows the customers to provide timely feedback to the supplier regarding how well its products or services satisfy their needs. This feedback mechanism from customers to the supplier should promote the supplier's innovation too. Second, a close supplier-customer relationship facilitates interpersonal interactions and helps employees (especially researchers)

[1] These effects include, for example, financing cost (Cen, Dasgupta, Elkamhi, and Pungaliya, 2014), capital structure decisions (Kale and Shahrur, 2007; Banerjee, Dasgupta, and Kim, 2008; and Chu, 2012), relationship-specific investments (Kale, Kedia, and Williams, 2011), cross-ownership (Fee, Hadlock, and Thomas, 2006), mergers and acquisitions (Fee and Thomas, 2004; Shahrur, 2005; and Ahern and Harford, 2014), and financial distress (Hertzel, Li, Officer, and Rodgers, 2008).

on both sides to share knowledge and exchange ideas on improving existing products and developing new products and technologies more efficiently, which helps enhance supplier innovation (Feldman, 1999; Audretsch and Feldman, 2004). Both arguments suggest that knowledge spillovers and timely feedback along the supply chain enhance supplier innovation, and they are supported by abundant anecdotal evidence observed in the economy. For example, Boeing, a large customer in our sample, interacts actively with its small suppliers and guide their research and development through a Mentor-Protégé Program. As Adex Machining Technologies, one of Boeing's small suppliers, describes:

> *"As a protégé, Adex is learning how to do business with Boeing, ... the learning process, which includes learning Boeing standards and procedures, is kind of like special forces training."*

In this paper, we aim to test this hypothesis – knowledge spillovers along the supplier chain enhance supplier innovation.

To tackle this research question, there are two major challenges. First, knowledge spillovers involve soft information production and transmission, which is difficult to directly observe and empirically capture. To overcome this hurdle, we use the geographical distance between a supplier and its major customers to capture knowledge spillovers along the supply chain. Although rapid development of transportation and communication tools in the last few decades has significantly reduced the cost of collecting hard information, acquiring soft information and facilitating knowledge spillovers through interpersonal interactions from a distance is still difficult and costly. Soft information is, by definition, different from hard information and is difficult to put down on paper, store electronically, or transfer to others (Petersen and Rajan, 2002). Collecting soft information and facilitating knowledge spillovers through frequent interpersonal interactions largely depends on the geographical distance between the parties involved in the supplier-customer relationship.[2] We therefore

[2]Many large customers rely on certain mentor programs to interact with their suppliers, which usually require frequent on-site visits and training. General Bearing Corp, one of small suppliers to Visteon in our

use a supplier's physical proximity to its customers to capture knowledge spillovers along the supply chain.

Second, identifying the casual effect of knowledge spillovers on firm innovation is challenging. The location choices of suppliers and customers are likely endogenous and affected by unobservable firm and market characteristics. Thus, a correlation between knowledge spillovers and supplier innovation may tell us little about the causal effect of knowledge spillovers on innovation. We overcome this identification challenge by exploiting plausibly exogenous variation in the geographical distance between a supplier and its major customers caused by customer relocation decisions in a generalized difference-in-differences framework and by undertaking a number of robustness analyses and placebo tests.

One important feature of the supplier-customer relationship based on Compustat segment customer database is that customers are much larger than their suppliers (i.e., more than 100 times larger in terms of total assets on average). This feature allows us to use customer firm headquarters relocations as plausible exogenous shocks to the geographical distance between the supplier and its customers, because arguably large customers are unlikely to change their headquarters in response to factors related to their suppliers that are much smaller than them.

Using a generalized difference-in-differences method, we find that the geographical distance between the supplier and its major customer has a negative effect on the quantity, quality, and efficiency of supplier innovation, which are measured by patent counts, number of citations per patent, and the ratio between patent counts and R&D investment accumulated (and depreciated) over the last five years, respectively. We verify that our baseline results are not driven by supplier's loss of business resulted from the termination of customer-

sample, proudly mentioned that "Visteon, one of our largest customers, has recognized us as an outstanding supplier and worthy of the support of their 'Lean Supplier Development' program. In November 2004, Mike Homan from Visteon visited our facility and conducted an assessment of our Lean activities. He did some additional training for the GBC Lean Team and made some suggestions for a Kaizen..... Mike returned to GBC in January 2005 to do more in depth training and lead us through a 5S Kaizen of three areas on the shop floor...... The event lasted 3 days and consisted of training, hands on exercises, and practical implementation of 5S principles."

supplier relationship after customer relocation. Actually, in the "moving-apart" relocation subsample in which customers move away from their suppliers, the relationship persists for more than 3 years after the relocation for all firms except for three cases, so the termination of relationship is unlikely to drive our results. Meanwhile, our results also hold in the subsample of "moving-closer" relocations, in which the termination of customer-supplier relationship is not a major concern.

To further establish causality, we address various concerns of our baseline identification strategy. First, while customers are much larger than their suppliers and hence customers are unlikely to relocate their headquarters simply for reasons related to the innovation of suppliers, we cannot completely rule out this possibility if we do not exactly observe customer relocation reasons. To address this concern, we manually search news for the exact reasons of customer relocations. We exclude customer relocations due to reasons that are related to suppliers and only include customer relocations that are categorized as for exogenous reasons. Examples of exogenous relocations include: move to retain or attract top executives, move to achieve low labor cost, move to take advantage of low real estate and living cost, move due to internal restructuring, mergers and acquisitions, and move closer to their own customers. Our main results are unchanged in the subsample in which customers relocate headquarters for exogenous reasons.

Second, one potential problem of our identification strategy is that customer relocation decisions could be correlated with local conditions that affect supplier innovation, which is not stated in their public announcements and hence cannot be captured by our test above. For example, customers may move to the city where the supplier locates because the city has favorable economic and social conditions, which can also positively affect supplier innovation. The same argument applies if customers move away from the city where the supplier locates in response to unfavorable economic and social conditions. To address this concern, we explicitly exclude customer relocations in which the customer is either moving to or moving away from the metropolitan areas where the supplier locates. We find that the results remain robust. To further address the possibility that the customer relocation decisions are correlated with local economic or social conditions, we add State times Year fixed effects in our baseline regressions. Including State times Year fixed effects can control any time-varying, confounding state level factors that can affect supplier innovation but are otherwise unobservable.

Third, because our baseline results hinge on the interaction between the supplier-customer pair, the documented effect should be absent if we artificially assign any two firms in a pair of supplier-customer relationship. We conduct two falsification tests to examine this conjecture. First, for each pair of supplier-customer in our sample, we fix suppliers and create an fictitious customer by finding a matched non-customer firm (based on 3-digit SIC industry classifications and firm assets) that best resembles the customer firm. We find that the effect of proximity between a supplier and its fictitious customer on its innovation is mixed and statistically insignificant. Second, for each pair of supplier-customer in our sample, we fix customers and create fictitious suppliers that are in the same state, the same 3-digit SIC industry classification, and have the closest assets as the true suppliers. Similarly, the proximity between a customer and its fictitious supplier has no effect on the supplier's innovation. Both falsification tests suggest that our baseline results are not driven by chance and are unlikely spurious.

Finally, there still remains a potential concern that an omitted variable coinciding with customer relocations could be the true underlying cause of changes in supplier innovation. If this is the case, then the changes in supplier innovation we attribute to customer headquarter relocations reflect mere an association rather than a causal effect. Our baseline identification strategy employs shocks (customer relocations) that affect different firms at different times. Hence, it is unlikely that an omitted variable unrelated to customer relocations would fluctuate every time (or even most of the times) customer relocation occurs. Therefore, our strategy of using multiple shocks due to customer relocations over time mitigates this concern. Still, we address this possibility by conducting another falsification test. Specifically, we begin by obtaining an empirical distribution of the relocation timing of customers in our sample. Next, we randomly assign the customer relocation timing (without replacement) to the customers who actually relocate their headquarters during our sample period. This approach maintains the distribution of customer relocation years from our baseline specification, but it disrupts the proper assignment of customer relocation years. Therefore, if an unobservable shock occurs at approximately the same time as the customer relocation years, it should still reside in the testing framework, and thus have an opportunity to drive the results. However, if no such shock exists, then our incorrect assignments of customer relocation years should weaken our results when we re-estimate the baseline tests. Indeed, we find these falsely assumed customer reallocations have no effect on innovation.

After demonstrating that there appears a positive, causal effect of knowledge spillovers from customers on supplier innovation, we explore possible underlying mechanisms through which knowledge spillovers affect firm innovation. We postulate that if knowledge spillovers between suppliers and customers are truly the driving force along the supply chain that affects supplier innovation, we expect the change in physical proximity to have a more pronounced effect on supplier innovation when the customer is more active in innovation activities and when both the customer and the supplier are in close technological proximity. Consistent with our conjecture, we find that the effect of proximity on supplier innovation is stronger when customers have higher R&D expenditures and a higher level of innovation output. We also find that the effect of the distance is stronger when the supplier and the customer are close in technological space.

In the final part of the paper, we explore a "bottom-line" question regarding the economic value of innovation due to knowledge spillovers from customers. We decompose a firm's innovation output into a component that is attributable to knowledge spillovers from customers and the other component that is unrelated to knowledge spillovers. We find that innovation attributable to knowledge spillovers from customers positively affects a firm's performance in product markets in terms of a higher return on assets (ROA), and this enhancement in ROA is driven by the improvement in asset turnover rate (i.e., sales/total assets), a performance metrics that is closely related to the supplier's customer needs, and is not driven by the change in profit margins (i.e., net income/sales), a performance metrics that is related to the supplier's own operating efficiency and hence is not related to satisfying its customer needs.

The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 describes the data and sample construction. Section 4 presents our main empirical results, and section 5 concludes.

2 Relation to the Existing Literature

This paper contributes to two strands of literature. First, our paper contributes to the growing literature on the interaction between supply chain relationships and corporate finance. One group of this literature examines how corporate financing and investments affect the supply-chain relationship. For example, Cen, Dasgupta, and Sen (2011) document that lower vulnerability to takeovers allows a supplier to establish a stable relationship with its customers, which encourages the supplier to make more customer-specific investment and results in better operating performance. Fee and Thomas (2004) and Shahrur (2005) find that mergers and acquisitions have negative effects on suppliers but insignificant effects on customers. Fee, Hadlock, and Thomas (2006) study the cross-ownership of firms along the supply chain and show that customer's equity ownership in the supplier mitigates hold-up problems and financial market frictions. Hertzel, Li, Officer, and Rodgers (2008) focus on the wealth effects of financial distress on suppliers and customers and find that firms in financial distress negatively affect their suppliers and customers.

The other group of literature instead examines how supplier-customer relationship may affect corporate financing. Banerjee, Dasgupta, and Kim (2008) find that the reliance on big customers or suppliers lowers a firm's leverage ratio. Kale and Shahrur (2007) find that when the supplier-customer relationship requires more relationship-specific investments, both the supplier and the customer maintain lower levels of leverage. Cen, Dasgupta, Elkamhi, and

Pungaliya (2014) show that a supplier's long-term relationships with principal customers has reputational consequences that reduce the supplier's cost of debt. Chu (2012) finds that intense supplier competition causes the firm to lower leverage. However, very little has been done on how supplier-customer relationships affect corporate real decisions. The only exception is Kale, Kedia, and Williams (2011), who study how CEO risk-taking incentives affect the incentive of customers and suppliers to engage in relationship-specific investments. Our paper also contributes to the literature by examining how knowledge spillovers along the supply chain affect corporate real decisions, more specifically, corporate innovation, an important real decision a company has to make to keep its competitive advantages.

Second, our paper contributes to the emerging literature on finance and innovation. This literature examines how various market and firm characteristics motivate and finance corporate innovation.[3] However, how knowledge spillovers along the supply chain and interactions between customers and suppliers affect a supplier's innovation is less well understood. Our paper is the first to tackle this research question. The supply chain aspect of enhancing innovation is important, because more and more firms outsource many of their innovation inputs to third party suppliers.

3 Data and Sample Construction

3.1 The sample

Our sample consists of all supplier-customer pairs that can be identified in Compustat between 1976 and 2009. We exclude utility firms (SIC code from 4900 to 4999) and financial firms (SIC code from 6000 to 6999) from our sample because these two industries are highly regulated. We also exclude non-innovative firms that file zero patents throughout our sample

[3] These factors include product market competition (Aghion, Harris, Howitt, and Vickers, 2001), bankruptcy laws (Acharya and Subramanian, 2009), labor laws and unions (Acharya, Baghai, and Subramanian, 2013; Acharya, Baghai, and Subramanian, 2014; and Bradley, Kim, and Tian, 2013), investor failure tolerance (Tian and Wang, 2014), stock liquidity (Fang, Tian, and Tice, 2014), firm boundaries (Seru, 2014), financial market development (Hsu, Tian, and Xu, 2014), analyst coverage (He and Tian, 2013), and banking competition (Cornaggia, Mao, Tian, and Wolfe, 2014).

period. According to the FASB 14 (1976) and 131 (1997), public firms are required to disclose customers who account for at least 10% of total sales, which allows us to identify major customers for a given firm.

A practical difficulty is that, while these disclosures are available in the Compustat segment files, the primary customers are only reported with abbreviated names without any other identifiers. To address this problem, we use a method similar to that of Fee and Thomas (2004) to match the reported customer names to Compustat firms. From the Compustat segment data file, we first exclude all of the customers that are reported as governments, regions, or militaries. We then run a text matching program to find the potential matches of the reported customer name with the Compustat firm names. The program requires all of the letters in the reported customer name to be sequentially presented in the potential match. To ensure matching accuracy, we manually identify customers from the matched pairs from the text matching program. If there are multiple potential matches and we cannot choose the unique match by screening the available public information (Firm web sites, annual reports, and Google), we conservatively exclude all these possible firm-customer pairs. Finally, we drop all pairs in which the reported customer is in the retail industry (SIC code 5200 to 5999), because retail customers are less likely to demand specific products and therefore are less likely to give valuable feedback that can help the suppliers improve their innovation. Our sample selection procedure results in a total of 8,645 firm customer pairs and 35,153 supplier-customer pair years. From the 35,153 pair year observations, we delete any observations for which the total assets or sales are either zero or negative and firm-year observations with missing data.

While the existing literature typically uses a firm's headquarters reported in Compustat to identify a firm's physical location, the Compustat location data only provides a snapshot of state and county information of firms' headquarters locations. This information is not sufficient to obtain the accurate information of corporate headquarter relocation, which we need for our analysis in this paper. To correct for this deficiency, we use Compact Disclosure, Corporate Library, and the Fortune Magazine to identify corporate headquarter relocations of customer firms. We are able to find 254 relocation cases, including 193 cases of cross-city relocations (44 of which are cross-state relocations) and 61 cases of within-city relocations. To capture meaningful change in distance, we focus on those cross-city relocations.[4] The cross-city relocations sample includes 2,933 firm-year observations, and 1,018 supplier-customer pairs with 869 unique suppliers and 120 unique customers. The relocations are not clustered in time. As shown in Table 2, the number of relocations is almost evenly distributed across time, and does not appear to exhibit strong correlation with business cycles or other economic conditions. The relocations are not clustered geographically either, so firms in our sample are not moving into or out of some specific areas.

We use the relocation data constructed above to test the effect of customers' knowledge spillovers on suppliers' innovation activities and outcomes in our baseline regression. A common concern of this identification strategy is that customer relocation decisions may be endogenous and possibly related to their suppliers. Therefore, it is important to understand the exact reasons for corporate relocations. Hence, we make a news search of Factiva, LexisNexis, and the Corporate Websites for the exact reasons of customer relocations. Among all the relocation cases, we are able to find relocation reasons for 45 cases. We summarize the relocation reasons into nine main categories in Table 2: (1) move close to customer, (2) move close to supplier, (3) retain or attract top executives, (4) low cost, (5) low real estate and living cost, (6) internal restructuring, (7) merger and acquisition related, (8) local government incentives, (9) reduce travel cost. Among these categories, only three categories — moving close to supplier, local government incentives and reducing travel cost — are potentially related to supplier unobservable characteristics. To address the potential concern of endogenous relocations, we exclude from our baseline regression the relocation cases that fall into these three categories and the relocation cases for which we cannot clearly identify the

[4]Since within-city relocations do not create meaningful change in distance, we use it as a natural falsification test reported in Panel A of Table 5. As expected, the within-city relocations which do not create much change in distance have no impact on supplier innovation.

underlying relocation reasons, and the results remain robust.

3.2 Variable measurement

3.2.1 Measuring innovation

We construct innovation variables using the NBER patent citation database initially created by Hall, Jaffe, and Trajtenberg (2001). This database provides detailed information on more than three million patents granted by the United States Patent and Trademark Office (USPTO) from 1976 to 2006. The patent database provides information on patent assignee names, 3-digit patent technology classes, and the number of future citations received by each patent. We then augment the NBER database with the Harvard Business School Patent Network Dataverse to extend the coverage to 2010.

Based on the augmented patent database, we construct two measures for innovation output. The first measure is the number of patent applications filed in a year that are eventually granted. This measure captures the quantity of innovation output. To capture the quality of innovation output, we construct a second measure by counting the total number of future citations a patent receives in subsequent years.

Following the existing literature, we adjust the output measures for two types of truncation problems. The first truncation problem arises as patents appear in the database only after they are granted and it may take several years for the USPTO to approve a patent. For example, if one firm files a patent application in 2009, and the patent is approved in 2011, the patent will not be included in our measure of patent output for 2009. To adjust this truncation bias, we follow Hall, Jaffe, and Trajtenberg (2001) to use the "weight factors" computed from the application-grant empirical distribution to adjust the patent counts. The second truncation problem arises as patents keep receiving citations over a long period, but we only observe the citations received up to 2010. We follow Hall, Jaffe, and Trajtenberg (2001) to adjust the truncation bias in citation counts by using the citation-lag distribution.

In addition to the two innovation output measures described above, we construct an innovation efficiency measure, which captures innovation output per unit of input, in which the innovation input is measured by R&D capital accumulated over the previous five years. Specifically, we follow Hirshleifer, Hsu, and Li (2013) to define accumulated R&D capital as the sum of R&D investment that is depreciated by an annual rate of 20% in the previous five years.

Finally, as shown in previous literature, the distribution of patent counts and citation counts is right skewed. We therefore use the natural logarithm of one plus the citation counts (*LnCites*), and one plus innovation efficiency (*LnIE*) as the innovation measures in our analysis.

3.2.2 Measuring distance and control variables

We calculate the distance variable as the geographical distance between the headquarters of the supplier and the headquarters of the customer. We collect information on historical headquarters addresses from Compact Disclosure and Fortune Magazine to augment the current headquarters address information in Compustat (Pirinsky and Wang (2006)). For each supplier and customer, we obtain the pair of latitude and longitude coordinates (measured in degrees of decimal) of its headquarters from the U.S. Census Bureau's Gazetteer City-State File. Because of the earth's near-spherical shape (technically an oblate spheroid), calculating an accurate distance between two points requires the use of spherical geometry and trigonometric math functions. We therefore convert latitude or longitude from decimal degrees to radians by dividing the latitude and longitude values by $180/n$, or approximately 57.296. Because the radius of the Earth is assumed to be 6,378.8 kilometers, or 3,963 miles, we use the Great Circle Distance Formula to calculate mileage between two pairs of latitudes and longitudes:

$$3963 \times \arccos[\sin(Lat_1) \times \sin(Lat_2) + \cos(Lat_1) \times \cos(Lat_2) \times \cos(Long_2 - Long_1)] \quad (1)$$

where Lat_1 and Lat_2 ($Long_1$ and $Long_2$) represent the latitudes (longitudes) of two points respectively. Because the distribution of distance is right skewed, we compute the natural logarithm of the distance (*LnDistance*) and use it as the main variable of interest.

We follow the existing literature to control for a vector of firm characteristics that may affect a firm's innovation output. The control variables include R&D Investment (R&D expenditure divided by total assets), Firm Size (natural logarithm of total assets), ROA (operating income divided by total assets), Tobin's Q (market value of assets divided by book value of total assets), Leverage (total debt divided by market value of assets), Sales Growth (growth rate of sales), Cash (cash holding divided by total assets), Tangibility (total property, plant, and equipment divided by total assets), Cap Exp (capital expenditures divided by total assets), Ln Age (natural logarithm of years listed in Compustat). In some specifications we also include customer characteristics, which are similarly defined as the supplier variables. All variable definitions are in Table 1.

3.3 Summary statistics

Table 3 provides summary statistics of the variables used in this study. An average supplier has about 13 patents a year, and each patent receives 9 future citations. These numbers are higher than those typically reported in previous innovation studies using Compustat firms for two possible reasons. First, we focus only on innovative suppliers, i.e., suppliers produced at least one patent over the sample period. Second, by sample construction, suppliers in our sample have large customers and are more likely to make relation-specific investment (Kale and Shahrur, 2007; Banerjee, Dasgupta, and Kim, 2008; Chu, 2012), which results in a higher level of innovation output.

The average distance between a supplier and its customer is 930 miles with a standard deviation of 890 miles. All other firm characteristics are comparable to those reported in the existing studies. Comparing the summary statistics of supplier variables with customer variables, one observation stands out —- customer firms are much larger than supplier firms, in fact they are about 123 times larger than supplier firms on average. This feature of the data is critical for our identification strategy used in this paper because these large customers are unlikely to change headquarters locations due to customer related factors given their suppliers are much smaller (and hence much less important) compared to them.

4 Empirical Results

In this section, we first discuss our baseline specification and present the baseline results. We find strong evidence showing the significant impact of customer-supplier distance on supplier innovation. We then address some potential concerns regarding our baseline identification strategy. Our baseline results hold steadily when we employ a more restricted subsample that is unlikely to suffer from the potential endogeneity problems. Various falsification tests also confirm that the impact of knowledge spillovers is customer-supplier pair specific, lending strong support to our baseline results.

The impact of customer-supplier distance on supplier innovation is found to be more pronounced for more innovative customers and for customers and suppliers that employ close technology, both of which suggest the knowledge spillovers along the supply chain as the overriding underlying mechanism.

We also find that the knowledge spillovers improve supplier's long-term product market performance (e.g., ROA) through boosting the supplier's innovation output, and this enhancement in the supplier's ROA is driven by the improvement in asset turnover which is a performance metrics closely related to customers rather than by changes in profit margins which mainly capture operation efficiency and cost of production and is not related to customer needs.

4.1 Baseline specifications and results

In this paper, we use the physical distance between suppliers and customers as a proxy of knowledge spillovers and investigate its effect on supplier's innovation output. However, the identification of the causal effect of knowledge spillovers on supplier's innovation is challenging, because geographical concentration and economic outcomes are often simultaneous determined. Specifically, in our setting, the location choices of suppliers or customers and the innovation activities could be simultaneous determined by some unobservables, leading to biased inferences from the standard Ordinary Least Square (OLS) regressions in which innovation measures are regressed on distance measures.

To overcome this hurdle and establish causality, our baseline identification strategy relies on one critical feature of the U.S. supplier-customer relationship observed in the Compustat segment customer database, i.e., customers are much larger than their suppliers (more than 100 times larger on average). Arguably, those large customers are unlikely to change their headquarters locations for reasons that are closely related to the innovation of their suppliers that are much smaller compared with them. Therefore, our baseline analysis uses a difference-in-differences approach that relies on the plausibly exogenous variation in distance driven by customer headquarters relocations for identification.

Specifically, we estimate the following model:

$$Innovation_{i\tau} = \alpha + \beta LnDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}, \qquad (2)$$

where i indexes firm, t indexes time, and j indexes industry. The dependent variable in this model is our measure of the supplier's innovation quantity (*LnPatents*), or quality (*LnCites*). X_{it} is a vector of supplier and customer characteristics. We include both the year fixed effects, $Year_t$, and supplier-customer pair fixed effects, $Pair_{ij}$, in our regression. This specification is a generalized difference-in-differences specification because the variation in $LnDistance_{ijt}$ only comes from the supplier-customer pairs in which customer headquarters relocation occurs. For supplier-customer pairs in which customers' headquarters locations remain unchanged in our sample period, $LnDistance_{ijt}$ is time-invariant.

Intuitively speaking, a short distance between supplier and customer facilitates face-to-face communication which could be very important for soft information production and transmission.[5] When the customer moves closer to the supplier for some arguably exogenous reasons, we expect them to have more efficient exchanges of ideas and knowledge, which provides timely feedback to suppliers about customers' needs and eventually increase the supplier's innovation output.

We report the regression results estimation in Table 4. Columns (1)-(3) show the regression results of innovation quantity measured as *LnPatents* in years $t+1$ to $t+3$. As expected, the coefficient estimates on *LnDistance* are all negative and statistically significant, suggesting a negative relation between the geographical distance between the supplier and its major customers and the supplier's future innovation patent counts. The economic effect is sizeable: One standard deviation increase in the distance from its mean leads to a 7% decrease in the number of patents filed next year. The results in column (2) and (3) suggest that the effects extend to patent filings in the next two and three years, respectively.

Columns (4)-(6) show the results for innovation quality measured as patent citations (*LnCites*). Since the dependent variable is only well defined if the supplier produces at lease one patent in the corresponding year, we therefore exclude all firm-year observations in which the supplier does not produce any patent. The coefficient estimates on *LnDistance* are again all negative and statistically significant in all three columns, suggesting that a long distance between a supplier and its major customers negatively affects the quality of its patents generated in the subsequent years. The effect is also economically large: one standard deviation increase in distance from its mean leads to a 12.5% decrease in the number of citations received per patent in the following year.

Lastly, columns (7)-(9) show the results for innovation efficiency, which is measured by

[5]See Uysal, Kedia, and Panchapagesan (2008) and Tian (2011) for a similar argument in the mergers and acquisition and venture capital investment settings, respectively.

the innovation output (patent) per unit of innovation input (R&D stock). We exclude all firm-year observations in which the supplier has zero total R&D expense over the last five years because the accumulated R&D expenses appears on the denominator of the innovation efficiency measure. The coefficient estimates on *LnDistance* are negative in all three columns and are statistically significant in columns (8) and (9). The evidence suggests that a firm's distance from its major customers negatively affects a firm's innovation efficiency, especially in the next 2 to 3 years.

Overall, our baseline results show that the distance between supplier and customer has significant impact on the supplier's innovation output. Suppliers' innovation quantity, quality, and efficiency all rise (drop) significantly after their major customers relocate closer to (further away from) them. The effect is persistent in the next three years following the relocation, confirming the long-lasting impact of knowledge spillovers on supplier's innovation activities and output.

A potential concern regarding our baseline results is that the impact of distance on supplier innovation does not truly capture the effect of knowledge spillovers but merely reflects supplier's loss of business due to the termination of customer-supplier relationship after the major customers move away. We rule out this competing explanation through two exercises. First, we check how long the existing customer-supplier relationships last after the major customers move away from the suppliers. We find that in only three cases the relationship terminates within three years of customer relocation, and the regression results remain quite similar even if we exclude these three cases in our regression. Still, one may argue that the relationship could simply become weaker even if not completely terminated after customers move away, which may also reduce suppliers' revenue and affect their innovation. We then exclude all "moving-apart" cases in which the major customers move away from their suppliers and rerun our regression using the remaining "moving-closer" cases in which customers move closer to their suppliers. Our baseline results hold steadily in this "moving-closer" subsample, which is unlikely to be driven by a weaker or terminated relationship.

4.2 Additional identification attempts

In this subsection, we undertake additional analyses in the difference-in-differences framework to address a few potential concerns regarding our main identification strategies.

We first show that our baseline results continue to hold when we restricted our analysis to a subsample in which the reasons of customer relocation can be clearly identified as exogenous. Next, we control for local economic conditions that can possibly create spurious correlation between the customer-supplier distance and suppliers' innovation, and the results become even stronger.

We then conduct three falsification tests to demonstrate that the knowledge spillover effect identified in the baseline analysis is truly specific to the observed customer-supplier pair to mitigate other endogeneity concerns that may arise from the omitted variables problem.

4.2.1 Addressing endogenous customer relocations

The key identification assumption in our baseline tests is that the customer's relocation decisions are uncorrelated with factors that may potentially affect a supplier's innovation activities. Though the large discrepancy in size between the customers and their suppliers helps mitigate this concern, we cannot completely rule out this possibility without knowing the exact reasons of customer relocation. We thus search through different sources such as Compact Disclosure, Corporate Library, and the Fortune Magazine to manually collect the reasons of corporate headquarter relocations of customer firms. Among all the relocation cases, we are able to find relocation reasons for 45 cases. We summarize the relocation reasons into nine main categories in Table 2: (1) move close to their own customers, (2) move close to the suppliers, (3) retain or attract top executives, (4) low cost, (5) low real estate and living cost, (6) internal restructuring, (7) merger and acquisition related, (8) local government incentives, (9) reduce travel cost. Among these categories, only three categories— moving close to supplier, local government incentives and reducing travel cost— are potentially related to supplier unobservable characteristics. We exclude the relocation cases falling into these three categories and the relocation cases for which we cannot clearly identify the reasons.

We then re-estimate Equation 2 in this restricted sample and report the results in Panel A of Table 5. Similar to Table 4, we report results for innovation quantity (*LnPatents*) in columns (1)-(3), innovation quality (*LnCites*) in columns (4)-(6), and innovation efficiency (*LnIE*) in column (7)-(9). The coefficient estimates on *LnDistance* are negative and significant at the 5% or 1% level in all regressions, and their magnitudes remain similar and economically sounded. This finding suggests that our baseline results are unlikely to be driven by customer relocation decisions that are correlated with supplier innovation activities.

One remaining concern is that even if we exclude customer relocations for stated reasons that are likely to be correlated with supplier innovation activities, customers may still move due to reasons that are not publicly stated but are related to supplier innovation. Local economic conditions could be such an unstated relocation reason. For example, the customer is in the same location as the supplier before relocation and then moves away from the current location due to unfavorable local economic conditions. Alternatively, the customer relocates to the same location as the supplier due to favorable local economic conditions. In the first case, unfavorable local economic conditions drive away the customer (and thus increase the distance between the supplier and the customer) and meanwhile decrease supplier innovation. In the second case, favorable local economic conditions attract the customer (and thus decrease the distance between the supplier and the customer) and meanwhile increase supplier innovation. Both cases may create spurious correlation between distance and supplier innovation. To explicitly address these concerns, we exclude customer headquarters relocations in which the customer either moves away from or moves to the same state as the supplier. We repeat our analysis by estimating equation 2 in this restricted sample and report the results in Panel B of Table 5.

The coefficient estimates on *LnDistance* are negative and statistically significant in all columns for innovation quantity, quality, and efficiency. In an untabulated analysis, we repeat the analysis in a sample in which we exclude customer headquarters relocations in which the customer either moves away from or moves to the same city as the supplier. We get even stronger results. Overall, our evidence suggests that the negative effect of the geographical distance between the supplier and its major customers documented in the baseline analysis is unlikely to be driven by local economic conditions that also affect customer relocation decisions.

To further address the possibility that the customer relocation decisions are correlated with local economic conditions, we add $State \times Year$ fixed effects to our baseline regressions. Including $State \times Year$ fixed effects can control any time-varying, confounding state level factors that can affect supplier innovation but are otherwise unobservable. The results with $State \times Year$ fixed effects controlled are presented in Panel C of Table 5. The coefficient estimates on *LnDistance* are very similar to those in Table 4, although we lose statistical significance in three out of nine specifications.

Overall, our baseline results continue to hold when we restrict our analysis to the subsample of relocations in which relocation reasons can be clearly identified to be exogenous to suppliers' innovation activities. The results are also robust to the inclusion of State times Year fixed effect, showing that our findings are unlikely to be driven by unobservable state-level factors.

4.2.2 Falsification tests

In this subsection, we conduct three sets of falsification tests to provide complementary evidence to support our main findings documented above.

First, in our regressions above, we exclude all within-city relocation cases and only keep the cross-city relocation cases in our sample. Our argument is that only cross-city relocations create meaningful change in the distance between suppliers and customers and are therefore expected to affect supplier innovation. This constructs the first set of falsification test for our baseline results. That is, our results should not hold for the subsample of within-city relocations because these relocation cases do not change the customer-supplier distance significantly. We rerun our baseline regression using the within-city relocation cases, and the results are reported in Panel A of Table 5. As expected, none of the coefficients are significant, indicating that it is the change in distance rather than the relocations per se that affects the supplier innovation.

Second, if knowledge spillovers from the major customers truly affect their suppliers' innovation, this effect has to take place through the specific customer-supplier pair. In other words, we shall not expect to observe any correlation between a firm's innovation output and its distance from another firm that is not its customer. We conduct the second set of falsification test to verify this, and the falsification test consists of two exercises. In the first exercise, for each customer-supplier pair observed in data, we take the customer as given and create a fictitious supplier for it. We select the fictitious supplier from the firms that are in the same state, in the same 3-digit SIC industry and have the closest total assets as the true supplier. The match is performed at the time when the true supplier and its customer first report their supplier-customer relationship, and we then follow the fictitious supplier-customer pair for the same number of years of the true supplier-customer relationship. We re-estimate Equation 2 with the fictitious suppliers. Because the fictitious supplier is in the same state as the true supplier, if the main results are driven by local economic conditions, we should still observe the effects on this falsification test. We report the results in Panel B of Table 5. In all columns, the coefficient estimates on *LnDistance* have mixed signs and none of them is statistically significant.

In the second exercise, we take the supplier as given and create a fictitious customer for it. The fictitious customer matches the true customer observed in data in the industry and total asset. We rerun the regressions estimating Equation 2 using the geographical distance between the supplier and this fictitious customer firm. We report the results in Panel C of Table 5. In all columns, the coefficient estimates on *LnDistance* have mixed signs and all of them are statistically insignificant. The two exercises suggest that our baseline results are absent in fictitiously assigned supplier-customer pairs, supporting the hypothesis that it is supplier-customer specific knowledge spillovers that drive our baseline results.

Finally, there still exists a potential concern that an omitted variable coinciding with customer relocations could be the true underlying cause of changes in supplier innovation. If this is the case, then the changes in supplier innovation we attribute to customer headquarter relocations reflect merely an associations rather than a causal effect. Our baseline identification strategy employs shocks (customer relocations) that affect different firms at different times. Hence, it is unlikely that an omitted variable unrelated to customer relocations would fluctuate every time (or even most of the times) customer relocation occurs. Therefore, our strategy of using multiple shocks due to customer relocations over time mitigates this concern. To further rule out this possibility, we conduct the third set of falsification test.

Specifically, we begin by obtaining an empirical distribution of the relocation timing of customers in our sample. Next, we randomly assign the customer relocation timing (without replacement) to the customers that actually relocate their headquarters during our sample period. This approach maintains the distribution of customer relocation years from our baseline specification, but it disrupts the proper assignment of customer relocation years. Therefore, if an unobservable shock occurs at approximately the same time as the customer relocation years, it should still reside in the testing framework, and thus have an opportunity to drive the results. However, if no such shock exists, then our incorrect assignments of customer relocation years should weaken our results when we re-estimate the baseline tests, because intuitively the changes in supplier innovation well before or well after the year of customer relocation should not be systematically correlated with the changes in distance occurred at the year of relocation.

We report the results in Panel D of Table 5. None of the coefficient estimates on *LnDistance* is statistically significant and the magnitudes of coefficient estimates are also small. These non-results corroborate the notion that our paper's main results are not driven by an omitted variable.

In addition to the falsification tests above, our results remain robust if we control for additional supplier and customer characteristics in the regressions. In fact, the magnitudes of the coefficients on *LnDistance* do not change much when we change the number of control variables. However, standard deviations do change when we increase or decrease the number of control variables, which further suggests that customer relocation decisions are likely exogenous (Roberts and Whited, 2012).

4.3 Possible mechanisms

In this subsection, we explore possible underlying economic mechanisms through which the geographical distance between the supplier and its major customers affects supplier innovation. If knowledge spillovers drive the results as we postulated, we should expect to observe significant cross-sectional variation in the results when the importance of knowledge spillovers varies across firms. In particular, we expect the results to be stronger if

(1) The customers are more innovative by themselves; Or

(2) The customers and suppliers employ closely related technologies.

(1) is intuitive as illustrated in a simple example: though both general retailers and auto producers could be big customers of tire producers, the feedbacks provided by auto producers will be more valuable in improving the tire producers' innovation than those provided by the general retailers. This argument is because auto producers know much better what improvement in tires will enhance the performance of autos given their own experiences in producing and improving autos.

The importance of (2) is motivated by Jaffe (1986), which shows that the effect of knowledge spillovers is stronger between firms that are close in technological space. In our context, if the distance affects supplier innovation through the knowledge spillover channel, the effect should be stronger if the supplier and the customer are close in technological space.

To test the first conjecture, we add two interaction terms in our baseline regressions: the interaction between *LnDistance* and customer R&D expenditures and the interaction between *LnDistance* and the number of patents the customer has. We believe that customers' R&D expenditure and their patents capture their own innovation intensity.

We present the results in Panel A of Table 7. The coefficient estimates on the interaction terms are negative in all columns and statistically significant mainly in regressions in which innovation efficiency is examined. Overall, these results suggest the effect of *LnDistance* on supplier innovation efficiency is stronger when the customers spend more on R&D or produce more innovation output. The evidence is consistent with the argument that knowledge spillovers from the customer to the supplier are an important channel through which the distance affects supplier innovation.

To test the second conjecture, we follow Jaffe (1986) to construct a measure for technology proximity between the supplier and the customer as follows:

$$TechnologyProximity = \frac{(S'C)^2}{(S'S)(C'C)}, \tag{3}$$

where S is a column vector, and each element of S is the ratio of the number of supplier's patents granted in the last three year in a patent class to the total number of supplier's patents granted in the last three years. The column vector C is similarly defined for customer's patents. The measure *Technology Proximity* is bounded between 0 and 1.

We then add the interaction term between *LnDistance* and *Technology Proximity* to our baseline regressions, and present the results in Panel B of Table 7. The coefficient estimates on the interaction term are all negative and are statistically significant in seven out of nine columns. Theses results suggest that the effect of *LnDistance* on supplier innovation is more pronounced if the supplier and the customer are closer in technological space. Together with the notion that technological proximity facilitates knowledge spillovers (Jaffe (1986)), the evidence is again consistent with the argument that knowledge spillovers from the customer to the supplier are an important channel through which the distance affects supplier innovation.

One implication of the knowledge spillovers mechanism is that a shorter distance between the supplier and its major customers should facilitate the interactions between them. It is, however, very difficult, if not impossible, to measure physical interactions between the supplier and its customer. We therefore focus on technological interactions between the supplier and its customer to examine whether shorter distance facilitates more technological interactions. To this end, we use cross-citations to measure technological interactions. Specifically, we use the natural logarithm of one plus the number of times that a supplier's patent cites its customer's patent (*LnCrossCitation*). We then run a regression similar to Equation 2, with the dependent variable replaced with *LnCrossCitation*.

We report the results in Panel C of Table 7. Consistent with our conjecture, the coefficient estimates on *LnDistance* are all negative and statistically significant, which suggests that a short distance between the supplier and its customer facilitates technological interactions between them, which positively contributes to the innovation output and efficiency of the supplier. The result is consistent with the argument the geographical distance affects supplier innovation through its effect on facilitating technological interactions between the supplier and its major customers.

4.4 Economic value implications

Finally, we explore a "bottom-line" question by examining whether the effect of knowledge spillovers on supplier innovation can ultimately improve supplier performance in the product market. Previous work has documented that increase in innovation has a positive effect on firms' performance such as firm valuation (Hall, Jaffe, and Trajtenberg, 2005). Thus, the main challenge we face is to identify the change in a supplier's innovation that could be attributed to the knowledge spillovers from customers. To overcome this hurdle, we first decompose the variation in a supplier's innovation into two components. The first component is the variation in innovation caused by knowledge spillovers from the customers; the second component is the residual variation in innovation that is orthogonal to the knowledge spillovers. Specifically, we first estimate the following model:

$$Innovation_{i\tau} = \alpha + \beta LnDistance_{ijt} + Year_t + Pair_{ij} + \varepsilon_{ijt}, \tag{4}$$

Note that the only difference between Equation 4 and Equation 2 is that Equation 4 does not include any control variables. With Equation 4, we can effectively estimate the part of innovation that can be attributed to customer knowledge spillovers. More formally, we denote the fitted value from the estimates of Equation 4 as *Patent Predicted*, and the residual from Equation 4 as *Patent Residual*, in which the dependent variable is *Ln Lead Patent*. By the nature of OLS regression, these two components are orthogonal. The first component is fully driven by knowledge spillovers, while the second component captures any remaining variation in innovation that is unrelated to the knowledge spillovers channel.

We focus on a firm's return on assets (ROA) as the main measure of the firm's performance, and regress a supplier's ROA on these two components — *Patent Predicted* and *Lead Patent Residual*. Since the key independent variables are generated variables, we use bootstrap to calculate the standard errors. Results are presented in Table 8. Columns (1)-(3) show that the coefficient estimates on *Lead Patent Predicted*, which capture the contribution of a supplier's innovation to its ROA that is related to the knowledge spillovers, are all positive and significant. One standard deviation increase of *Lead Patent Predicted* (0.26) increases ROA by about 0.03, which is quite sizable. The coefficient estimates on *Lead Patent Residual*, however, are all close to zero in magnitude and insignificant statistically.

A closer look at the ROA results yields some even more interesting findings. Specifically, we explore the driving force of an improvement in a firm's ROA. This analysis is feasible because following the DuPont analysis, we can decompose ROA into the product of two components: asset turnover ($\frac{Sales}{Total\,Assets}$) and profit margins ($\frac{Net\,Income}{Sales}$). Asset turnover measures a firm's revenue generating ability and is a metrics directly related to how good a supplier can satisfy its customer needs. However, profit margins captures a firm's operating efficiency and is a metrics related to the firm's own operating cost and its own suppliers, which is not directly related to the satisfaction of the firm's customers. If knowledge spillovers along the supply chain promotes innovation, which in turn enhances the firm's ability to satisfy their customers' needs, it should improve the firm's asset turnover but should have no effect on its profit margin. We investigate the impact of *Patent Predicted* and *Lead Patent Residual* defined above on these two performance metrics by regressing asset turnover and profit margins on *Patent Predicted* and *Lead Patent Residual*. Columns (4)-(9) in Table 8 show that supplier's innovation driven by knowledge spillovers from the customers significantly affects the supplier's asset turnover, but has negligible impact on the supplier's profit margins. Therefore, our findings here highlight the knowledge spillovers as the main driving force of the supplier's product market performance improvement.

5 Conclusion

In this paper, we examine the effect of supplier-customer relationship on supplier innovation through a knowledge spillovers channel. We use the geographical distance between a supplier and its major customers to capture knowledge spillovers. To establish causality, we explore plausibly exogenous variation in distance caused by customer headquarters relocations. In a generalized difference-in-differences framework, we show that knowledge spillovers from customers have a positive, causal effect on supplier innovation. Our finding is consistent with the argument that knowledge spillovers facilitated by feedback provided by customers and frequent interactions with customers enhance supplier innovation. We also find that the effect of knowledge spillovers on supplier innovation is stronger when the customers are more R&D intensive and are more innovative themselves and when the customers are in closer technology proximity with the suppliers. Finally, we show that innovation attributable to knowledge spillovers from customers positively contributes to a firm's product market performance, and further more, the improvement in product market performance is mainly driven by the enhancement in the customer-related performance metrics. Our paper provides new insights into the real effect of knowledge spillovers along the supply chain and its enhancement on firm innovation.

References

Acharya, V. V., R. P. Baghai, K. V. Subramanian, 2013. Labor laws and innovation. Journal of Law and Economics 56(4), 997–1037.

——— , 2014. Wrongful discharge laws and innovation. Review of Financial Studies 27(1), 301–346.

Acharya, V. V., K. V. Subramanian, 2009. Bankruptcy codes and innovation. Review of Financial Studies 22(12), 4949–4988.

Aghion, P., C. Harris, P. Howitt, J. Vickers, 2001. Competition, imitation and growth with step-by-step innovation. Review of Economic Studies 68(3), 467–492.

Ahern, K., J. Harford, 2014. The importance of industry links in merger waves. Journal of Finance 69(2), 527–576.

Audretsch, D. B., M. P. Feldman, 2004. Knowledge spillovers and the geography of innovation. Handbook of Regional and Urban Economics 4, 2713–2739.

Banerjee, S., S. Dasgupta, Y. Kim, 2008. Buyer–supplier relationships and the stakeholder theory of capital structure. Journal of Finance 63(5), 2507–2552.

Bradley, D., I. Kim, X. Tian, 2013. The causal effect of labor unions on innovation. Working paper.

Cen, L., S. Dasgupta, R. Elkamhi, R. S. Pungaliya, 2014. Reputation and loan contract terms: the role of principal customers. Working Paper, University of Toronto.

Cen, L., S. Dasgupta, R. Sen, 2011. Discipline or disruption? stakeholder relationships and the effect of takeover threat. Working Paper, Hong Kong University of Science and Technology.

Chu, Y., 2012. Optimal capital structure,bargaining, and the supplier market structure. Journal of Financial Economics 106(2), 411–426.

Cornaggia, J., Y. Mao, X. Tian, B. Wolfe, 2014. Does banking competition affect innovation?. Journal of Financial Economics, Forthcoming.

Fang, V. W., X. Tian, S. Tice, 2014. Does stock liquidity enhance or impede Firm Innovation?. Journal of Finance 69(5), 2085–2125.

Fee, C. E., C. J. Hadlock, S. Thomas, 2006. Corporate equity ownership and the governance of product market relationships. Journal of Finance 61(3), 1217–1251.

Fee, C. E., S. Thomas, 2004. Sources of gains in horizontal mergers: evidence from customer, supplier, and rival firms. Journal of Financial Economics 74(3), 423–460.

Feldman, M. P., 1999. The new economics of innovation, spillovers and agglomeration: A review of empirical studies. Economics of Innovation and New Technology 8(1-2), 5–25.

Hall, B. H., A. Jaffe, M. Trajtenberg, 2005. Market Value and Patent Citations. RAND Journal of Economics 36(1), 16–38.

Hall, B. H., A. B. Jaffe, M. Trajtenberg, 2001. The NBER patent citation data file: Lessons, insights and methodological tools. Unpublished working paper. National Bureau of Economic Research.

Han, J. K., N. Kim, R. Srivastava, 1998. Market Orientation and Organizational Performance: Is innovation a Missing Link?. Journal of Marketing 62(4), 30–45.

He, J. J., X. Tian, 2013. The dark side of analyst coverage: The case of innovation. Journal of Financial Economics 109(3), 856–878.

Hertzel, M. G., Z. Li, M. S. Officer, K. J. Rodgers, 2008. Inter-firm linkages and the wealth effects of financial distress along the supply chain. Journal of Financial Economics 87(2), 374–387.

Hirshleifer, D., P.-H. Hsu, D. Li, 2013. Innovative efficiency and stock returns. Journal of Financial Economics 107(3), 632–654.

Hsu, P.-H., X. Tian, Y. Xu, 2014. Financial development and innovation: Cross-country evidence. Journal of Financial Economics 112(1), 116–135.

Jaffe, A., 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. American Economic Review 76(5), 984–1001.

Kale, J., S. Kedia, R. Williams, 2011. The effect of CEO risk-taking incentives on relationship-specific investments by customers and suppliers. Working Paper, Northeastern University.

Kale, J., H. Shahrur, 2007. Corporate capital structure and the characteristics of suppliers and customers. Journal of Financial Economics 83(2), 321–365.

Lukas, B., O. Ferrell, 2000. The effect of market orientation on product innovation. Journal of the Academy of Marketing Science 28(2), 239–247.

Manso, G., 2011. Motivating innovation. Journal of Finance 66(5), 1823–1869.

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

Pirinsky, C., Q. Wang, 2006. Does corporate headquarters location matter for stock returns?. Journal of Finance 61(4), 1991–2015.

Porter, M., 1992. Capital disadvantage: America's failing capital investment system. Harvard Business Review 19(4), 65–82.

Roberts, M., T. Whited, 2012. Endogeneity in empirical corporate finance. Handbook of the Economics of Finance 2.

Seru, A., 2014. Firm boundaries matter: Evidence from conglomerates and R&D activity. Journal of Financial Economics 111(2), 381–405.

Shahrur, H., 2005. Industry structure and horizontal takeovers: Analysis of wealth effects on rivals, suppliers, and corporate customers. Journal of Financial Economics 76(1), 61–98.

Tian, X., 2011. The causes and consequences of venture capital stage financing. Journal of Financial Economics 101(1), 132–159.

Tian, X., T. Y. Wang, 2014. Tolerance for failure and corporate innovation. Review of Financial Studies 27(1), 211–255.

Uysal, V. B., S. Kedia, V. Panchapagesan, 2008. Geography and acquirer returns. Journal of Financial Intermediation 17(2), 256–275.

Table 1: Variable Definitions

Variable	Definition
LnPatents	Natural logarithm of one plus the number of patents filed (and eventually granted) of the supplier
LnCites	Natural logarithm of one plus the number of citations received on the supplier's patents filed (and eventually granted)
LnIE	Natural logarithm of one plus the ratio of number of patents to accumulated R&D expense (xrd+0.8 xrd (t-1)+0.6 xrd (t-2) +0.4 xrd (t-3)+0.2 xrd (t-4))
LnDistance	Natural logarithm of the geographical distance between the headquarters of the supplier and its customer
Technology Proximity	$\frac{(S'C)^2}{(S'S)(C'C)}$, where S and C are vectors of the ratios of patents awarded in patent classes to total patents for suppliers and customers
R&D	R&D expense divided by total assets
Q	Market value of total assets to book value of total assets
ROA	Net income divided by total assets
Leverage	Book value of total debt divided by market value of total assets
Log Assets	Natural logarithm of total assets
Sale Growth	The growth rate of sales
Cash	Cash holding divided by total assets
Tangibility	Total property, plant, and equipment divided by total assets
Cap Ex	Capital expenditure divided by total assets
Ln Age	Natural logarithm of the number of years in Compustat
Asset Turnover	Sales divided by total assets
Profit Margins	Net income divided by total sales

Table 2: The Distribution of Customer Relocations

The table reports the distribution of customer relocations in different years and for different reasons. The relocations are identified by searching Compact Disclosure, Corporate Library, and the Fortune Magazine. The reasons of relocations are identified by news searching of Factiva, LexisNexis, and the Corporate Websites.

Years	Number of Relocations	Moving Reason	Number of Relocations
1976-1979	5	Close to customer	2
1980-1984	28	Close to supplier	1
1985-1989	32	Retain or attract top executives	2
1990-1994	31	Low cost	12
1995-1999	53	Low real estate or living cost	2
2000-2004	28	Internal restructuring	15
2005-2009	16	M&A related	9
		Local government incentive	1
		Reduce travel cost	1
		Unknown	148

Table 3: Summary Statistics

This table reports the summary statistics for variables used in this paper. *Patent* is the number of patents filed (and eventually granted), *Cite* is the number of citations received on the patents filed, *Innovation Efficiency* is the ratio of number of patents to accumulated R&D expense (xrd+0.8 xrd (t-1)+0.6 xrd (t-2) +0.4 xrd (t-3)+0.2 xrd (t-4)), Q is market value of total assets to book value of total assets, *R&D* is R&D expense divided by total assets, *ROA* is the operating income divided by total assets, *Leverage* is the book value of total debt divided by market value of total assets, *Sales Growth* is the growth rate of sales, *Cash* is the cash holding divided by total assets, *Tangibility* is total property, plant, and equipment divided by total assets, *Cap Ex* is the capital expenditure divided by total assets, *Ln Age* is the natural logarithm of the number of years in Compustat, *Distance* is the geographical distance (in miles) between the headquarters of the supplier and its customer, and *Technology Proximity* is computed as $\frac{(S'C)^2}{(S'S)(C'C)}$, where S and C are vectors of the ratios of patents awarded in patent classes to total patents for suppliers and customers

Variable	obs	Mean	Std. Dev.	p25	Median	p75
Supplier						
Patent	8,333	13.94	98.53	0.00	1.00	4.00
Cite	8,333	8.85	19.52	0.00	0.00	11.25
Innovation Efficiency	7,438	0.38	12.08	0.00	0.01	0.13
Q	8,333	1.90	1.97	0.80	1.20	2.22
R&D	8,333	0.10	0.14	0.01	0.05	0.13
ROA	8,333	0.04	0.24	0.00	0.11	0.17
Leverage	8,333	0.20	0.23	0.00	0.11	0.33
Ln Assets	8,333	5.06	1.95	3.68	5.00	6.37
Sales Growth	8,333	0.26	1.00	-0.05	0.10	0.30
Cash	8,333	0.24	0.24	0.04	0.16	0.38
Tangibility	8,333	0.24	0.17	0.09	0.20	0.35
Cap EX	8,333	0.06	0.06	0.02	0.04	0.08
Ln Age	8,333	2.36	0.68	1.95	2.40	2.89
Customer						
Patent	8,333	334.83	587.83	4.00	121.00	423.00
Cite	8,333	12.97	12.10	1.16	12.66	18.09
Innovation Efficiency	7,442	0.05	2.22	0.00	0.00	0.00
Q	7,312	1.47	1.43	0.64	0.92	1.72
R&D	7,610	0.05	0.04	0.02	0.05	0.07
ROA	8,325	0.14	0.08	0.09	0.13	0.19
Leverage	7,312	0.29	0.27	0.08	0.20	0.40
Ln Assets	8,333	9.85	1.77	8.74	10.10	11.06
Sales Growth	8,295	0.11	0.20	0.01	0.08	0.17
Cash	8,332	0.13	0.12	0.05	0.09	0.17
Tangibility	8,333	0.24	0.15	0.11	0.22	0.34
Cap EX	8,300	0.06	0.05	0.03	0.05	0.09
Ln Age	8,314	2.81	0.59	2.48	2.89	3.26
Supplier-Customer Pair						
Distance	8,333	939	891	168	588	1658
Technology Proximity	8,499	0.03	0.10	0.00	0.00	0.00

Table 4: Baseline Regression Results

This table reports the baseline regression results of the model $Innovation_{i\tau} = \alpha + \beta LnDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variables are *LnPatents* in Columns (1)-(3), *LnCites* in Columns (4)-(6), and *LnIE* in Columns (7)-(9). Definitions of variables are listed in Table 1. Year fixed effects and supplier-customer pair fixed effects are included in all regressions. Robust standard errors are reported in parentheses below coefficient estimates. Significance levels at 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.072**	-0.051*	-0.040**	-0.126***	-0.043*	-0.211***	-0.059	-0.073**	-0.126***
	(0.028)	(0.027)	(0.019)	(0.036)	(0.025)	(0.027)	(0.058)	(0.034)	(0.046)
Q	0.002	0.005	0.015**	-0.012	0.035***	0.011	-0.007	-0.006	0.010
	(0.007)	(0.007)	(0.007)	(0.014)	(0.009)	(0.012)	(0.014)	(0.017)	(0.013)
R&D	0.453***	0.390*	-0.068	0.337	0.568*	0.153			
	(0.174)	(0.221)	(0.223)	(0.370)	(0.338)	(0.496)			
ROA	0.054	0.124	0.017	-0.013	0.212	0.290	0.262	0.071	0.118
	(0.096)	(0.115)	(0.105)	(0.244)	(0.261)	(0.270)	(0.233)	(0.265)	(0.227)
Leverage	-0.221	-0.281*	-0.254*	-0.206	0.025	0.115	-0.496**	-0.327	0.077
	(0.183)	(0.159)	(0.153)	(0.186)	(0.195)	(0.187)	(0.243)	(0.258)	(0.308)
Ln Assets	0.305***	0.242***	0.163**	-0.011	0.050	0.005	-0.211***	-0.154	-0.130
	(0.048)	(0.057)	(0.064)	(0.082)	(0.086)	(0.084)	(0.079)	(0.095)	(0.110)
Sale Growth	0.000	-0.000	-0.000	-0.000*	0.001	-0.000***	0.002	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.000)
Cash	0.130	-0.028	0.003	0.108	-0.147	0.201	0.268	0.006	0.108
	(0.176)	(0.205)	(0.225)	(0.244)	(0.342)	(0.301)	(0.279)	(0.350)	(0.357)
Tangibility	0.346	0.008	0.048	-0.356	-1.020*	-0.229	1.009	0.741	-0.065
	(0.400)	(0.415)	(0.418)	(0.607)	(0.613)	(0.654)	(0.972)	(1.067)	(0.910)
Cap EX	-0.364	-0.266	-0.528	0.229	0.080	-0.171	-0.688	-1.596**	-0.529
	(0.373)	(0.380)	(0.390)	(0.691)	(0.693)	(0.735)	(0.703)	(0.792)	(0.878)
Ln Age	0.376**	0.344*	0.264	-0.592***	-0.231	-0.317	0.220	0.170	0.113
	(0.170)	(0.209)	(0.214)	(0.219)	(0.222)	(0.223)	(0.385)	(0.412)	(0.320)
Customer R&D	0.099	-0.247	0.239	0.026	-0.100	-0.407	0.122	-0.920	0.204
	(0.337)	(0.490)	(0.612)	(0.765)	(0.618)	(1.027)	(0.743)	(0.687)	(1.052)
Customer Ln Assets	-0.103*	-0.052	-0.063	0.040	0.040	0.070	-0.058	-0.144	-0.221**
	(0.060)	(0.067)	(0.074)	(0.086)	(0.084)	(0.083)	(0.116)	(0.109)	(0.108)
Constant	1.014	1.068*	1.432**	2.524***	1.751*	2.641**	1.412	1.262	1.741
	(0.679)	(0.624)	(0.700)	(0.762)	(1.000)	(1.172)	(1.083)	(1.303)	(1.235)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,000	6,700	6,386	3,725	3,392	3,131	3,461	3,148	2,872
R-squared	0.856	0.846	0.845	0.791	0.790	0.794	0.865	0.862	0.871

Table 5: Addressing the potential endogeneity of customer relocation decisions

This table reports four sets of tests aimed at addressing the potential bias caused the endogeneity of customer relocation decisions. Panel A reports the regression results of the model in Equation 2 excluding customer relocations that are categorized as being related to the suppliers. Panel B reports the regression results of the model in Equation 2 excluding customer relocations in which the customer is either moving to the same state as the supplier or moving away from the same state as the supplier. Panel C reports the regression results with state/year fixed effects. The dependent variables are *LnPatents* in Columns (1)-(3), *LnCites* in Columns (4)-(6), and *LnIE* in Columns (7)-(9). Control variables are the same as in Table 4, but are omitted for brevity. Relevant control variables, year fixed effects and supplier-customer pair fixed effects are included in all regressions. Robust standard errors are reported in parentheses below coefficient estimates. Significance levels at 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

Panel A: Excluding relocations related to supplier and for unknown reasons

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.058***	-0.034***	-0.034***	-0.111***	-0.035**	-0.210***	-0.026**	-0.064***	-0.142***
	(0.009)	(0.010)	(0.011)	(0.011)	(0.015)	(0.022)	(0.011)	(0.016)	(0.022)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,359	6,098	5,823	3,400	3,104	2,857	3,153	2,880	2,621
R-squared	0.854	0.846	0.845	0.793	0.791	0.795	0.876	0.870	0.877

Panel B: Excluding customer relocating to or away from the same state as the supplier

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.051***	-0.034***	-0.034***	-0.120***	-0.048***	-0.221***	-0.017*	-0.056***	-0.129***
	(0.008)	(0.009)	(0.010)	(0.012)	(0.015)	(0.019)	(0.010)	(0.015)	(0.021)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,927	5,699	5,459	3,175	2,903	2,682	2,948	2,700	2,471
R-squared	0.849	0.839	0.840	0.789	0.790	0.793	0.870	0.864	0.869

Panel C: Results with state/year fixed effects

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.079*	-0.052*	-0.040	-0.100***	-0.016	-0.255***	-0.066*	-0.074	-0.213***
	(0.044)	(0.031)	(0.028)	(0.032)	(0.046)	(0.037)	(0.038)	(0.047)	(0.040)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,036	6,743	6,432	3,723	3,383	3,124	3,453	3,133	2,860
R-squared	0.889	0.882	0.880	0.865	0.865	0.874	0.922	0.923	0.925

Table 6: Falsification tests

This table reports four falsification tests. Panel A reports the falsification test results of the model $Innovation_{i\tau} = \alpha + \beta LnDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$ when only within-city relocations are included. Panel B reports the falsification test results of the model with fictitiously assigned suppliers and Panel C reports the falsification test results with fictitiously assigned customers. The fictitious supplier or customer is in the same three-digit industry as the true supplier or customer and is closest in firm size. Panel C reports the falsification test results of the model $Innovation_{i\tau} = \alpha + \beta LnDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$ with randomized relocation timing. The dependent variables are *LnPatents* in Columns (1)-(3), *LnCites* in Columns (4)-(6), and *LnIE* in Columns (7)-(9). Definitions of variables are listed in Table 1. Relevant control variables, year fixed effects and supplier-customer pair fixed effects are included in all regressions. Robust standard errors are reported in parentheses below coefficient estimates. Significance levels at 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

Panel A: Falsification tests with within-city relocations

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.007	-0.005	0.011	0.002	0.001	0.001	0.004	0.006	-0.002
	(0.097)	(0.121)	(0.135)	(0.181)	(0.218)	(0.246)	(0.172)	(0.176)	(0.185)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,984	6,312	5,666	3,689	3,158	2,763	3,419	2,917	2,527
R-squared	0.857	0.843	0.836	0.792	0.802	0.801	0.872	0.868	0.875

Panel B: Falsification tests with fictitiously assigned matched suppliers

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.029	-0.028	-0.019	-0.027*	0.007	0.012	0.005	-0.011	0.011
	(0.020)	(0.021)	(0.024)	(0.016)	(0.018)	(0.017)	(0.028)	(0.025)	(0.025)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,151	4,932	4,749	3,011	2,733	2,522	2,736	2,485	2,278
R-squared	0.740	0.716	0.709	0.794	0.782	0.791	0.779	0.795	0.793

Panel C: Falsification tests with fictitiously assigned matched customers

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	0.004	0.004	-0.005	-0.007	0.006	-0.008	0.006	0.001	-0.007
	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.016)	(0.012)	(0.012)	(0.011)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,482	6,177	5,886	3,493	3,189	2,976	3,237	2,952	2,728
R-squared	0.865	0.857	0.856	0.792	0.795	0.788	0.872	0.865	0.870

Panel D: Falsification tests with randomized relocation timing

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	0.023	0.021	-0.008	0.038*	0.022	-0.026	-0.024	0.003	-0.015
	(0.044)	(0.038)	(0.050)	(0.022)	(0.039)	(0.030)	(0.036)	(0.029)	(0.044)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,994	6,694	6,380	3,723	3,389	3,128	3,459	3,145	2,869
R-squared	0.856	0.846	0.845	0.791	0.789	0.793	0.865	0.862	0.871

Table 7: The mechanisms

This table reports regression results of the tests for possible mechanisms of the negative effects of distance on supplier innovation. Panel A reports the regression results of the model $Innovation_{i\tau} = \alpha + \beta_1 LnDistance_{ijt} + \beta_2 * LnDistance * LnCustomerPatent + \beta_3 * LnDistance * CustomerR\&D + \gamma_1 X_{it} + \gamma_2 Y_{jt} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variables are *LnPatents* in Columns (1)-(3), *LnCites* in Columns (4)-(6), and *LnIE* in Columns (7)-(9). Two interaction terms between *LnDistance* and *Ln Customer Patent*, *Customer R&D* are included in the regressions. Definitions of variables are listed in Table 1. Panel B reports the regression results of the model $Innovation_{i\tau} = \alpha + \beta_1 LnDistance_{ijt} + \beta_2 * LnDistance * TechnologyProximity + \gamma_1 X_{it} + \gamma_2 Y_{jt} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variables are *LnPatents* in Columns (1)-(3), *LnCites* in Columns (4)-(6), and *LnIE* in Columns (7)-(9). The interaction term between *LnDistance* and *Technology Proximity* is included in the regressions. Panel C reports the regression results of the model $Innovation_{i\tau} = \alpha + \beta LnDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variable is *LnCrossCitation*, which is defined as the number of times a supplier's patent cites its customer's patent. Relevant control variables, year fixed effects and supplier-customer pair fixed effects are included in all regressions. Robust standard errors are reported in parentheses below coefficient estimates. Significance levels at 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

Panel A: The effects of customer R&D expense and patents

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.001	0.007	0.040	-0.060	-0.030	-0.179**	0.075	0.084	0.083
	(0.055)	(0.059)	(0.063)	(0.080)	(0.081)	(0.091)	(0.084)	(0.084)	(0.092)
LnDistance ×	-0.012*	-0.010	-0.014*	-0.011	-0.003	-0.007	-0.018*	-0.024**	-0.033***
Ln Customer Patent	(0.007)	(0.008)	(0.008)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)
LnDistance ×	-0.106	-0.059	-0.130	-0.467	-0.402	-0.494	-1.267***	-1.030***	-1.123***
Customer R&D	(0.160)	(0.226)	(0.233)	(0.299)	(0.352)	(0.365)	(0.322)	(0.371)	(0.376)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,000	6,700	6,386	3,725	3,392	3,131	3,479	3,168	2,893
R-squared	0.856	0.847	0.845	0.791	0.790	0.795	0.866	0.862	0.873

Panel B: The effect of technological proximity

	LnPatents			*LnCites*			*LnIE*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LnDistance	-0.083**	-0.059*	-0.044	-0.123	-0.042	-0.219***	-0.073	-0.078	-0.127***
	(0.034)	(0.032)	(0.038)	(0.115)	(0.187)	(0.062)	(0.062)	(0.068)	(0.032)
LnDistance X	-0.012	-0.036*	-0.066***	-0.096*	-0.130*	-0.043	-0.058**	-0.057**	-0.058**
Technology Proximity	(0.021)	(0.020)	(0.021)	(0.056)	(0.068)	(0.076)	(0.027)	(0.028)	(0.027)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,000	6,700	6,386	3,725	3,392	3,131	3,461	3,148	2,872
R-squared	0.856	0.847	0.845	0.792	0.790	0.794	0.866	0.862	0.871

Panel C: The number of times a supplier's patent cites its customer's patent

Dependent Variable	LnCrossCitation		
	t+1	t+2	t+3
	(1)	(2)	(3)
LnDistance	-0.027***	-0.027**	-0.031*
	(0.009)	(0.011)	(0.017)
Control Variables	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	6,895	5,056	3,661
R-squared	0.573	0.583	0.580

Table 8: The outcome

This table reports the regression results of *ROA* and its two components, *Asset Turnover* and *Profit Margins*, on predicted patents and the residuals, in which the predicted patents and the residuals are calculated by regressing *LnPatents* on *LnDistance*, year fixed effects, and pair fixed effects. *ROA* is the return on total assets, defined as net income divided by total assets. Following DuPont analysis, *ROA* is decomposed into *Asset Turnover* and *Profit Margins*, where *Asset Turnover* is defined as total sales divided by the total assets, and *Profit Margins* is defined as net income divided by total sales. The following control variables are included: *Ln Assets*, *Ln Assets*, *Tangibility*, *Cap Ex*, *R&D*, and *Ln Age*. For brevity, the coefficient estimates on control variables are not reported. Bootstrap standard errors are reported in parentheses below coefficient estimates. Significance levels at 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

	ROA			*Asset Turnover*			*Profit Margins*		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lead Patent Predicted	0.118**	0.128**	0.137**	0.112***	0.114***	0.120***	-0.160	-0.055	0.014
	(0.050)	(0.059)	(0.065)	(0.026)	(0.027)	(0.028)	(0.147)	(0.227)	(0.384)
Lead Patent Residual	0.003	0.003	0.007	-0.020*	-0.020*	-0.023**	0.003	-0.091	-0.035
	(0.004)	(0.004)	(0.009)	(0.010)	(0.011)	(0.011)	(0.032)	(0.063)	(0.051)
Ln Assets	-0.039***	-0.025*	-0.019	-0.069***	-0.068***	-0.069***	0.056	0.018	0.105
	(0.008)	(0.013)	(0.015)	(0.005)	(0.005)	(0.005)	(0.043)	(0.091)	(0.108)
Tangibility	0.110**	0.025	-0.080	0.201***	0.223***	0.215***	0.641**	0.768	0.146
	(0.048)	(0.065)	(0.085)	(0.067)	(0.068)	(0.069)	(0.290)	(0.497)	(0.423)
Cap EX	-0.062	0.054	0.225*	-0.362**	-0.389***	-0.403***	0.172	-0.703	0.753
	(0.064)	(0.077)	(0.116)	(0.147)	(0.148)	(0.150)	(0.471)	(0.676)	(0.576)
Ln Age	0.012	-0.019	-0.028	0.131***	0.132***	0.131***	-0.060	0.093	-0.160
	(0.016)	(0.021)	(0.031)	(0.009)	(0.009)	(0.009)	(0.090)	(0.192)	(0.204)
R&D Expense	-0.062	0.046	0.043	-0.255***	-0.239***	-0.275***	0.456	-0.098	-1.231
	(0.083)	(0.118)	(0.136)	(0.073)	(0.077)	(0.077)	(0.909)	(1.106)	(1.300)
Constant	0.082	0.094	-0.006	1.428***	1.383***	1.385***	-0.092	-0.488	-0.603
	(0.161)	(0.173)	(0.178)	(0.109)	(0.126)	(0.125)	(0.209)	(0.385)	(1.500)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,806	7,383	6,829	7,765	7,477	7,176	6,018	5,261	4,500
R-squared	0.766	0.737	0.698	0.710	0.712	0.711	0.807	0.718	0.718