Learning from Customers: Corporate Innovation along the Supply Chain

Yongqiang Chu, Xuan Tian, and Wenyu Wang*

Current Version: March, 2014

Abstract

This paper studies 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 along the supply chain. 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 customer feedbacks and suppliers' frequent interpersonal interactions with customers enhance supplier innovation. We further show that the effect is stronger when the customers are more R&D intensive and are more innovative themselves. 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

1 Introduction

A growing literature has examined various effects of supplier-customer relationship on corporate finance decisions, such as capital structure decisions (Kale and Shahrur, 2007; Banerjee, Dasgupta, and Kim, 2008; and Chu, 2012), cross-ownership (Fee, Hadlock, and Thomas, 2006), mergers and acquisitions (Fee and Thomas, 2004; Shahrur, 2005; Ahern, 2012; and Ahern and Harford, 2013), and financial distress (Hertzel, Li, Officer, and Rodgers, 2008). While existing studies have highlighted the importance of the interactions between suppliers and customers along the supply chain in corporate finance, these studies have focused on how supplier-customer relationships affect corporate 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 investments – technological innovation, which is critical for a firm's long-term competitive advantages and sustainable growth (Porter, 1992), and a key underlying channel through which supplier-customer relationship affects innovation – knowledge spillovers.

Supplier-customer relationships could affect corporate innovation through knowledge spillovers in several different ways. First, a close relationship between a supplier and its major customers enables the supplier to learn the specific needs of its customers, and therefore stimulates more research and development (R&D) spending that ultimately leads to technological innovation on the supplier side (Han, Kim, and Srivastava, 1998; Lukas and Ferrell, 2000). Second, a close supplier-customer relationship helps employees especially researchers on both sides to share knowledge and exchange ideas more efficiently, which helps enhance supplier innovation (Feldman, 1999; Audretsch and Feldman, 2004). Both suggest that knowledge spillovers between the supplier and its customers enhance supplier innovation. In this paper, we focus on how knowledge spillovers along the supply chain affect corporate 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. We therefore use a supplier's 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 difference-in-differences framework.

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.

To further establish causality, we address various concerns of our baseline identification strategy. First, one potential problem of our identification strategy is that customer relocation decisions could be correlated with local conditions that affect supplier innovation. 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.

Second, while customers are much larger than their suppliers and hence it is unlikely that customer 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 moving 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.[1] We obtain similar and even stronger results.

Third, because our baseline results hinges on the interaction between the supplier-customer pair, if our baseline results are not spurious, the documented effect should be absent if we artificially assign any two firms in a pair of supplier-customer relationship. We conduct a falsification test to test this conjecture. For each pair of supplier-customer in our sample, we create an artificial customer by finding a matched non-customer firm (based on

[1]Examples of exogenous relocations include: move close to their own customers, 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.

3-digit SIC industry classifications and firm assets) that best resembles the customer firm. We find that the effect of a supplier's proximity to a randomly assigned customer on its innovation is mixed and statistically insignificant, which suggests that our baseline results are not driven by chance and spurious.

Finally, there still exists a remaining 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 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. 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 affect supplier innovation, we expect the change in geographical distance to have a more significant effect on supplier innovation when the customer is more active in innovation activities. We find that the effect of distance on supplier innovation is stronger when customers have higher R&D expenditures and higher innovation output. We also postulate that if knowledge spillovers are truly the driving force along the supply chain that affects supplier innovation, proximity should positively attribute to the technological interactions between the supplier and its customers. We use cross-citations to capture technological interaction and find that distance has a negative effect on the number of customers' patents being cited by the supplier's patents.

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 Related Literature

This paper contributes to three strands of literature. First, our paper contributes to the growing literature on the interaction between supply chain relationships and corporate finance. Fee and Thomas (2004) 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 customer's equity ownership in the supplier mitigates hold up problem and financial market frictions. Shahrur (2005) examine the effect of leverage on relationship-specific investments between suppliers and customers. The results suggest firms choose leverage to incentivize their customers and suppliers to make relationship-specific investments. 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 negative affect their suppliers and customers. Chu (2012) examines the effect of supplier competition on customer capital structure and finds that intense supplier competition causes the firm to lower leverage. In a recent paper, Chu, Huang, and Zhang (2014) find that asymmetric information between the supplier and its supplier can mitigate the holdup problem.

Second, our paper contributes to the emerging literature on finance and innovation. This literature examines the effect of various market and firm factors on corporate innovation, such as competition (Aghion, Harris, Howitt, and Vickers, 2001), bankruptcy laws (Acharya and Subramanian, 2009), labor laws and labor union (Acharya, Baghai, and Subramanian, forthcoming; Acharya, Baghai, and Subramanian, 2014; and Bradley, Kim, and Tian, 2013), investor failure tolerance (Tian and Wang, 2014), stock liquidity (Fang, Tian, and Tice, 2013), 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, 2013). Our paper is the first to explore the effects of knowledge spillovers along the supply chain on corporate innovation. The supply chain aspect of innovation can potentially be important as more and more firms outsource many of their inputs to third party suppliers.

Finally, our paper is related to the literature on knowledge spillovers. Most early literature on the relation between innovation and knowledge spillovers extends the knowledge production function framework of Griliches (1979) and Pakes and Griliches (1984) by adding a spatial element to the knowledge production function, and then estimates the augmented knowledge production function at the geographic unit level. Jaffe (1989) finds that firms innovate more if other firms or academic research institutions in the same area are also more innovative. Later studies largely followed Jaffe (1989)'s methodology and test the theories with refined innovation output and input measures (e.g., Acs, Audretsch, and Feldman, 1992; Anselin, Varga, and Acs, 1997; Autant-Bernard, 2001b; Autant-Bernard, 2001a; and Black, 2005).

While most of these studies find evidence consistent with knowledge spillovers, they suffer from two inherent problems. First, these studies are unable to identify the sources and beneficiaries of knowledge spillovers and therefore the exact mechanisms of knowledge spillovers. Our paper, by focusing on supplier-customer pairs, is able to explicitly identify knowledge spillovers between two firms and therefore is able to shed more light on the underlying mechanism of knowledge spillovers. Second, most existing studies suffer from the inherent endogeneity of location choices. Our paper addresses the identification issue by relying on plausibly exogenous variation caused by customer headquarters relocation.

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 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 and 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). Our sample selection procedure results in a total of 8,645 firm customer pairs and 35,153 firm customer years. From the 35,153 firm 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, one problem with the Compustat location data is that it only provides a snapshot of state and county information of firms' headquarters locations. 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, 44 of which are cross-state relocations and 193 of which are cross-city relocations. To capture meaningful relocations, we focus on those cross-city relocations. 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.

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 (*Log Cites*), and one plus innovation efficiency (*Log IE*) 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 (*Log Distance*) 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 Ex (capital expenditures divided by total assets), Log Age (Log 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 2 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, suppliers with reported customers often make relation-specific investments (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

4.1 Baseline specifications and results

Identify the causal effect knowledge spillovers measured by the physical distance between the supplier and its major customers on its own innovation output is challenging, because geographical concentration and economic outcomes are often simultaneous determined. Specifically, in our setting, the location choices of suppliers or customers and innovation activities could be simultaneous determined by unboservables that may seriously bias our estimation. Therefore, it is very difficult to make causal 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 identification strategy relies on one critical feature of the U.S. supplier-customer relationship observed from 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:

$$Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}, \tag{2}$$

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 *Log Patents*, or quality *Log Cites*. 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 $LogDistance_{ijt}$ 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, $LogDistance_{ijt}$ is time-invariant.

We report the regression results estimating in Table 3. Panel A shows the regression results of innovation quantity measured as *Log Patents* in years $t+1$ to $t+3$. The coefficient estimates on *Log Distance* are all negative and statistically significant, suggesting a negative relation between the geographical distance between the supplier and its major customers on the supplier's subsequent innovation patent counts. The economic effect is sizeable: One standard deviation increase in distance from its mean leads to a 7% decrease in patents filed in one year. The results in Column (2) and (3) suggest that the effects extend to patent filings in two and three years, respectively.

Panel B shows the results for innovation quality measured as *Log Cites*. Since the dependent variable is only well defined if the supplier produces at lease one patent in a year, we therefore exclude all firm-year observations in which the supplier does not produce any patent. The coefficient estimates on *Log Distance* are again all negative and statistically significant in all three columns, suggesting that a long distance between a firm and its major customers negatively affects the quality of its patents generated in subsequent years. The economic effect of distance is 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, Panel C shows the results for innovation efficiency. 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 *Log Distance* are negative in all three columns and are statistically significant in columns (2) and (3). The evidence suggests that a firm's proximity to its major customers negative affects a firm'ss innovation efficiency.

4.2 Addressing identification challenges

In this subsection, we undertake additional analyses in the difference-in-differences framework to address a few concerns on our main identification strategies, which could contaminate our inferences from the baseline analysis.

First, 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. However, this identification assumption could be potentially violated if: (1) 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; (2) 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 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 in this restricted sample.

We report the results in Table 4. Similar to Table 3, we report results for innovation quantity (*Log Patents*) in Panel A, innovation quality (*Log Cites*) in Panel B, and innovation efficiency (*Log IE*) in Panel C. The coefficient estimates on *Log Distance* are negative and statistically significant. 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.

Second, another concern regarding our identification is that although customers are much larger than their suppliers and hence it is unlikely that they relocate their headquarters for reasons related to the innovation of suppliers, we cannot completely rule out this possibility if we do not exactly observe their moving reasons. Hence, we make a news search for the exact reasons of customer relocations. Among all the relocation cases, we are able to find relocation reasons for 43 cases. We then categorize the moving reasons into eight categories:

(1) close to customer, (2) close to supplier, (3) retain or attract top executives, (4) low labor cost, (4) low real estate and living cost, (5) internal restructuring, (6) merger and acquisition related, (7) local government incentives, (8) reduce travel cost. Among these categories, only two categories, close to supplier and reduce travel cost, are potentially related to supplier unobservable characteristics. We therefore exclude observations of customer relocations that fall into these two categories and observations of customer relocations for which we cannot identify their exact relocation reasons. We then re-estimate Equation 2 in this restricted sample.

We report the results in Table 5. The coefficient estimates on *Log Distance* are negative and statistically significant in all regressions. This finding suggests that our baseline results are unlikely to be driven by customer relocation decisions that are correlated with supplier innovation activities.

Third, our baseline results hinges on the interaction between the supplier-customer pair, which means that one should not expect to observe our main results if we artificially assign any two firms in a pair of supplier-customer relationship. We conduct a falsification test to confirm this conjecture. For each pair of supplier-customer in our sample, we create an artificial customer by finding a matched non-customer firm that best resembles the customer firm (we match firms by in 3-digit SIC industry and book value of firm assets). We then assign this matched firm to the original supplier to create our artificial supplier-customer pair in the falsification test. We rerun the regressions estimating Equation 2 using the geographical distance between the supplier and this matched customer firm.

We report the results in Table 6. In all three panels, the coefficient estimates on *Log Distance* have mixed signs and none of them is statistically insignificant. This falsification test result suggests that our baseline results are absent in artificially assigned supplier-customer pairs and supplier-customer specific knowledge spillovers drive our baseline results.

Finally, there 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 mere 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.

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.

We report the results in Table 7. None of the coefficient estimates on *Log Distance* 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 above tests, our results remain robust if we control for additional supplier and customer characteristics in the regressions. In fact, the magnitudes of the coefficients on *Log Distance* 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, the effect should be stronger if the customer is also innovation intensive and the innovation activities of supplier and customer are more connected. To test this conjecture, we add two interaction terms in our baseline regressions: the interaction between *Log Distance* and customer R&D expenditures and the interaction between *Log Distance* and the number of patents the customer has.

We present the results in Table 8. The coefficient estimates on the interaction terms are all negative, but are statistically significant mainly in Panel C in which the supplier's innovation efficiency is examined. Overall, these results suggest the effect of *Log Distance* 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.

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 a customer's patent (*Log Cross Citation*). We then run a regression similar to Equation 2, with the dependent variable replaced with *Log Cross Citation*. We report the results in Table 9. Consistent with our conjecture, the coefficient estimates on *Log Distance* 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.

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. Our paper provides new insight onto the real effect of knowledge spillovers along the supply chain its enhancement on firm innovation.

REFERENCES

Acharya, V. V., R. P. Baghai, K. V. Subramanian, 2014. Wrongful discharge laws and innovation. Review of Financial Studies 27(1), 301–346.

——— , forthcoming. Labor laws and innovation. Journal of Law and Economics.

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

Acs, Z. J., D. B. Audretsch, M. P. Feldman, 1992. Real Effects of Academic Research: Comment. The American Economic Review 82(1), pp. 363–367.

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., 2012. Bargaining Power and Industry Dependence in Mergers. Journal of Financial Economics 103, 530–550.

Ahern, K., J. Harford, 2013. The Importance of Industry Links in Merger Waves. Journal of Finance, Forthcoming.

Anselin, L., A. Varga, Z. Acs, 1997. Local geographic spillovers between university research and high technology innovations. Journal of Urban Economics 42(3), 422–448.

Audretsch, D. B., M. P. Feldman, 2004. Knowledge spillovers and the geography of innovation. Handbook of regional and urban economics 4, 2713–2739.

Autant-Bernard, C., 2001a. The geography of knowledge spillovers and technological proximity. Economics of Innovation and New Technology 10(4), 237–254.

——— , 2001b. Science and knowledge flows: evidence from the French case. Research Policy 30(7), 1069–1078.

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

Black, G., 2005. The geography of small firm innovation. Springer.

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

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

Chu, Y., R. Huang, D. Zhang, 2014. Asymmetric Information and Holdup: Evidence from Supplier-Customer Relationships.. Working Paper, University of South Carolina.

Cornaggia, J., Y. Mao, X. Tian, B. Wolfe, 2013. Does Banking Competition Affect Innovation?. Journal of Financial Economics, Forthcoming.

Fang, V. W., X. Tian, S. Tice, 2013. Does Stock Liquidity Enhance or Impede Firm Innovation?. Journal of Finance, Forthcoming.

Fee, C. E., C. J. Hadlock, S. Thomas, 2006. Corporate equity ownership and the governance of product market relationships. The 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.

Griliches, Z., 1979. Issues in assessing the contribution of research and development to productivity growth. Bell Journal of economics 10(1), 92–116.

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, 116–135.

Jaffe, A. B., 1989. Real effects of academic research. The American Economic Review 79, 957–970.

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. Academy of Marketing Science 28, 239–247.

Pakes, A., Z. Griliches, 1984. Patents and R&D at the firm level: a first look. in *R & D, patents, and productivity*. University of Chicago Press pp. 55–72.

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

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

Porter, M., 1992. Capital disadvantage: Americas 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, 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., T. Y. Wang, 2014. Tolerance for failure and corporate innovation. Review of Financial Studies 27(1), 211–255.

Table 1: Variable Definitions

Variable	Definition
Log Patents	Natural logarithm of one plus the number of patents filed (and eventually granted) of the supplier
Log Cites	Natural logarithm of one plus the number of citations received on the supplier's patents filed (and eventually granted)
Log IE	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))
Log Distance	Natural logarithm of the geographical distance between the headquarters of the supplier and its customer
R&D	R&D expense divided by total assets
Q	Market value of total assets to book value of total assets
ROA	Operating income divided by total assets
Leverage	Book value of total debt divided by market value 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
Log Age	Natural logarithm of the number of years in Compustat

Table 2: Summary Statistics

This table reports the summary statistics for variables used in this paper. The variables are defined as in Table 1.

Variable	obs	Mean	Std. Dev.	p25	Median	p75
Patent	11784	12.64	90.32	0.00	1.00	4.00
Cite	11784	9.06	20.53	0.00	0.00	11.14
Innovation Efficiency	10136	0.25	2.33	0.00	0.01	0.13
Distance	10041	0.93	0.89	0.16	0.58	1.65
Q	10067	2.07	3.46	0.80	1.23	2.28
R&D	11784	0.12	0.22	0.01	0.06	0.15
ROA	11747	0.02	0.49	-0.01	0.10	0.17
Leverage	10067	0.19	0.23	0.00	0.10	0.32
Log Assets	11763	4.97	2.05	3.55	4.88	6.36
Sales Growth	11104	13.24	499.82	-0.05	0.10	0.31
Cash	11762	0.25	0.25	0.04	0.16	0.40
Tangibility	11760	0.24	0.17	0.09	0.20	0.35
Cap EX	11703	0.06	0.06	0.02	0.04	0.08
Log Age	11783	2.15	0.81	1.61	2.20	2.77
Customer R&D	11784	0.05	0.06	0.02	0.04	0.07
Customer Patent	11469	326.29	596.77	4.00	116.00	406.00
Customer Log Assets	11443	9.79	1.79	8.75	10.07	10.96
Customer Q	9847	1.58	2.08	0.65	0.94	1.80
Customer Leverage	9847	0.28	0.27	0.08	0.19	0.40
Customer ROA	11429	0.14	0.13	0.09	0.13	0.19
Customer Tangibility	11443	0.24	0.15	0.11	0.22	0.33
Customer Sales Growth	11247	191.12	12875.44	0.01	0.09	0.18
Customer Cash	11441	0.13	0.12	0.05	0.09	0.17

Table 3: Baseline Regression Results

This table reports the baseline regression results of the model $Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variables are *Log Patents* in Panel A, *Log Cites* in Panel B, and *Log IE* in Panel C. 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.

Panel A: Number of Patents

	(1)	(2)	(3)
Dependent Variable	*Log Patents (t+1)*	*Log Patents (t+2)*	*Log Patents (t+3)*
Log Distance	-0.072**	-0.051*	-0.040**
	(0.028)	(0.027)	(0.019)
Q	0.002	0.005	0.015**
	(0.007)	(0.007)	(0.007)
R&D	0.453***	0.390*	-0.068
	(0.174)	(0.221)	(0.223)
ROA	0.054	0.124	0.017
	(0.096)	(0.115)	(0.105)
Leverage	-0.221	-0.281*	-0.254*
	(0.183)	(0.159)	(0.153)
Log Assets	0.305***	0.242***	0.163**
	(0.048)	(0.057)	(0.064)
Sale Growth	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Cash	0.130	-0.028	0.003
	(0.176)	(0.205)	(0.225)
Tangibility	0.346	0.008	0.048
	(0.400)	(0.415)	(0.418)
Cap EX	-0.364	-0.266	-0.528
	(0.373)	(0.380)	(0.390)
Log Age	0.376**	0.344*	0.264
	(0.170)	(0.209)	(0.214)
Customer R&D	0.099	-0.247	0.239
	(0.337)	(0.490)	(0.612)
Customer Log Assets	-0.103*	-0.052	-0.063
	(0.060)	(0.067)	(0.074)
Constant	1.014	1.068*	1.432**
	(0.679)	(0.624)	(0.700)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	7,000	6,700	6,386
R-squared	0.856	0.846	0.845

Panel B: Number of Citations per Patent

	(1)	(2)	(3)
Dependent Variable	*Log Cites (t+1)*	*Log Cites (t+2)*	*Log Cites (t+3)*
Log Distance	-0.126***	-0.043*	-0.211***
	(0.036)	(0.025)	(0.027)
Q	-0.012	0.035***	0.011
	(0.014)	(0.009)	(0.012)
R&D	0.337	0.568*	0.153
	(0.370)	(0.338)	(0.496)
ROA	-0.013	0.212	0.290
	(0.244)	(0.261)	(0.270)
Leverage	-0.206	0.025	0.115
	(0.186)	(0.195)	(0.187)
Log Assets	-0.011	0.050	0.005
	(0.082)	(0.086)	(0.084)
Sale Growth	-0.000*	0.001	-0.000***
	(0.000)	(0.001)	(0.000)
Cash	0.108	-0.147	0.201
	(0.244)	(0.342)	(0.301)
Tangibility	-0.356	-1.020*	-0.229
	(0.607)	(0.613)	(0.654)
Cap EX	0.229	0.080	-0.171
	(0.691)	(0.693)	(0.735)
Log Age	-0.592***	-0.231	-0.317
	(0.219)	(0.222)	(0.223)
Customer R&D	0.026	-0.100	-0.407
	(0.765)	(0.618)	(1.027)
Customer Log Assets	0.040	0.040	0.070
	(0.086)	(0.084)	(0.083)
Constant	2.524***	1.751*	2.641**
	(0.762)	(1.000)	(1.172)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,725	3,392	3,131
R-squared	0.791	0.790	0.794

Panel C: Innovation Efficiency

	(1)	(2)	(3)
Dependent Variable	*Log IE(t+1)*	*Log IE (t+2)*	*Log IE(t+3)*
Log Distance	-0.059	-0.073**	-0.126***
	(0.058)	(0.034)	(0.046)
Q	-0.007	-0.006	0.010
	(0.014)	(0.017)	(0.013)
ROA	0.262	0.071	0.118
	(0.233)	(0.265)	(0.227)
Leverage	-0.496**	-0.327	0.077
	(0.243)	(0.258)	(0.308)
Log Assets	-0.211***	-0.154	-0.130
	(0.079)	(0.095)	(0.110)
Sale Growth	0.002	0.000	0.000
	(0.002)	(0.001)	(0.000)
Cash	0.268	0.006	0.108
	(0.279)	(0.350)	(0.357)
Tangibility	1.009	0.741	-0.065
	(0.972)	(1.067)	(0.910)
Cap EX	-0.688	-1.596**	-0.529
	(0.703)	(0.792)	(0.878)
Log Age	0.220	0.170	0.113
	(0.385)	(0.412)	(0.320)
Customer R&D	0.122	-0.920	0.204
	(0.743)	(0.687)	(1.052)
Customer Log Assets	-0.058	-0.144	-0.221**
	(0.116)	(0.109)	(0.108)
Constant	1.412	1.262	1.741
	(1.083)	(1.303)	(1.235)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,461	3,148	2,872
R-squared	0.865	0.862	0.871

Table 4: Regression results excluding customer relocating to or away from the same state as the supplier

This table reports the regression results of the model $Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$, 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. The dependent variables are *Log Patents* in Panel A, *Log Cites* in Panel B, and *Log IE* in Panel C. 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.

Panel A: Number of Patents

	(1)	(2)	(3)
Dependent Variable	*Log Patents (t+1)*	*Log Patents (t+2)*	*Log Patents (t+3)*
Log Distance	-0.418*	-0.419*	-0.217
	(0.226)	(0.243)	(0.238)
Q	0.000	0.010	0.015**
	(0.007)	(0.007)	(0.007)
R&D	0.432**	0.377*	-0.126
	(0.178)	(0.227)	(0.225)
ROA	0.074	0.174	0.028
	(0.102)	(0.123)	(0.110)
Leverage	-0.225	-0.259*	-0.266*
	(0.190)	(0.157)	(0.149)
Log Assets	0.309***	0.256***	0.172**
	(0.053)	(0.062)	(0.068)
Sale Growth	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Cash	0.157	-0.069	0.001
	(0.186)	(0.204)	(0.232)
Tangibility	0.377	-0.000	-0.022
	(0.405)	(0.412)	(0.414)
Cap EX	-0.393	-0.257	-0.523
	(0.404)	(0.389)	(0.403)
Log Age	0.346**	0.291	0.226
	(0.174)	(0.210)	(0.212)
Customer R&D	0.027**	0.018	0.019
	(0.013)	(0.014)	(0.013)
Customer Log Assets	-0.141**	-0.063	-0.089
	(0.066)	(0.075)	(0.083)
2008.fyear	-2.179***		
	(0.765)		
Constant	3.298**	3.187**	2.692*
	(1.504)	(1.581)	(1.607)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	6,193	5,932	5,657
R-squared	0.854	0.842	0.841

Panel B: Number of Citations per Patent

	(1)	(2)	(3)
Dependent Variable	*Log Cites (t+1)*	*Log Cites (t+2)*	*Log Cites (t+3)*
Log Distance	-0.120***	-0.048***	-0.221***
	(0.012)	(0.015)	(0.019)
Q	-0.005	0.035***	0.013
	(0.015)	(0.009)	(0.014)
R&D	0.214	0.580*	0.251
	(0.380)	(0.336)	(0.515)
ROA	-0.106	0.173	0.352
	(0.268)	(0.269)	(0.296)
Leverage	-0.270	0.003	0.146
	(0.204)	(0.206)	(0.198)
Log Assets	-0.056	0.039	0.006
	(0.082)	(0.090)	(0.084)
Sale Growth	-0.000	0.001	-0.000***
	(0.000)	(0.001)	(0.000)
Cash	0.110	0.060	0.295
	(0.257)	(0.343)	(0.324)
Tangibility	-0.208	-0.863	-0.191
	(0.662)	(0.652)	(0.683)
Cap EX	0.364	0.234	-0.245
	(0.703)	(0.711)	(0.726)
Log Age	-0.545**	-0.198	-0.239
	(0.235)	(0.232)	(0.223)
Customer R&D	-0.073	-0.017	0.002
	(0.084)	(0.091)	(0.037)
Customer Log Assets	0.049	0.037	0.061
	(0.087)	(0.093)	(0.093)
Constant	2.600***	1.580	2.673**
	(0.752)	(1.057)	(1.237)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,175	2,903	2,682
R-squared	0.789	0.790	0.793

Panel C: Innovation Efficiency

	(1)	(2)	(3)
Dependent Variable	*Log IE (t+1)*	*Log IE (t+2)*	*Log IE (t+3)*
Log Distance	-0.017*	-0.056***	-0.129***
	(0.010)	(0.015)	(0.021)
Q	-0.004	-0.006	0.017
	(0.015)	(0.019)	(0.012)
ROA	0.189	0.027	-0.001
	(0.236)	(0.280)	(0.243)
Leverage	-0.566**	-0.438	-0.024
	(0.248)	(0.273)	(0.324)
Log Assets	-0.198**	-0.109	-0.091
	(0.092)	(0.107)	(0.119)
Sale Growth	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.000)
Cash	0.270	0.200	0.124
	(0.306)	(0.397)	(0.403)
Tangibility	1.000	0.845	-0.103
	(0.991)	(1.119)	(0.913)
Cap EX	-0.471	-1.331	-0.289
	(0.743)	(0.832)	(0.915)
Log Age	0.396	0.235	0.139
	(0.407)	(0.440)	(0.342)
Customer R&D	0.057	0.083	0.007
	(0.058)	(0.051)	(0.036)
Customer Log Assets	0.008	-0.096	-0.187
	(0.117)	(0.123)	(0.123)
Constant	0.849	0.643	1.496
	(1.091)	(1.419)	(1.347)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	2,948	2,700	2,471
R-squared	0.870	0.864	0.869

Table 5: Regression results excluding relocations related to supplier and for unknown reasons
This table reports the regression results of the model $Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$, excluding customer relocations for reasons potentially related to their suppliers or unknown reasons. The dependent variables are *Log Patents* in Panel A, *Log Cites* in Panel B, and *Log IE* in Panel C. 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.

Panel A: Number of Patents

	(1)	(2)	(3)
Dependent Variable	*Log Patents (t+1)*	*Log Patents (t+2)*	*Log Patents (t+3)*
Log Distance	-0.058***	-0.034***	-0.034***
	(0.009)	(0.010)	(0.011)
Q	0.000	0.007	0.014**
	(0.007)	(0.006)	(0.007)
R&D	0.439**	0.365	-0.062
	(0.180)	(0.230)	(0.220)
ROA	0.075	0.146	0.040
	(0.099)	(0.121)	(0.105)
Leverage	-0.234	-0.301*	-0.294*
	(0.194)	(0.167)	(0.160)
Log Assets	0.318***	0.268***	0.187***
	(0.052)	(0.060)	(0.068)
Sale Growth	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Cash	0.152	-0.049	-0.036
	(0.183)	(0.202)	(0.234)
Tangibility	0.401	0.046	0.052
	(0.434)	(0.445)	(0.445)
Cap EX	-0.352	-0.313	-0.648
	(0.405)	(0.402)	(0.419)
Log Age	0.389**	0.372*	0.295
	(0.180)	(0.222)	(0.226)
Customer R&D	0.028**	0.019	0.020*
	(0.013)	(0.014)	(0.012)
Customer Log Assets	-0.115*	-0.063	-0.080
	(0.064)	(0.072)	(0.080)
Constant	0.981	0.957	1.513**
	(0.687)	(0.638)	(0.736)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	6,359	6,098	5,823
R-squared	0.854	0.846	0.845

Panel B: Number of Citations per Patent

	(1)	(2)	(3)
Dependent Variable	*Log Cites (t+1)*	*Log Cites (t+2)*	*Log Cites (t+3)*
Log Distance	-0.111***	-0.035**	-0.210***
	(0.011)	(0.015)	(0.022)
Q	-0.020	0.031***	0.008
	(0.014)	(0.010)	(0.016)
R&D	0.348	0.639*	0.206
	(0.378)	(0.342)	(0.516)
ROA	-0.036	0.254	0.321
	(0.254)	(0.263)	(0.286)
Leverage	-0.211	0.025	0.108
	(0.193)	(0.201)	(0.192)
Log Assets	-0.058	0.042	0.013
	(0.081)	(0.090)	(0.085)
Sale Growth	-0.000	0.001	-0.001***
	(0.000)	(0.001)	(0.000)
Cash	0.167	-0.016	0.284
	(0.263)	(0.362)	(0.320)
Tangibility	-0.498	-1.149*	-0.231
	(0.625)	(0.642)	(0.682)
Cap EX	0.450	0.085	-0.191
	(0.710)	(0.723)	(0.756)
Log Age	-0.555**	-0.221	-0.274
	(0.228)	(0.231)	(0.227)
Customer R&D	-0.072	-0.015	0.004
	(0.090)	(0.091)	(0.037)
Customer Log Assets	0.084	0.061	0.093
	(0.087)	(0.087)	(0.084)
Constant	2.332***	1.577	2.388**
	(0.761)	(1.012)	(1.189)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,400	3,104	2,857
R-squared	0.793	0.791	0.795

Panel C: Innovation Efficiency

	(1)	(2)	(3)
Dependent Variable	*Log IE (t+1)*	*Log IE (t+2)*	*Log IE (t+3)*
Log Distance	-0.026**	-0.064***	-0.142***
	(0.011)	(0.016)	(0.022)
Q	-0.001	0.006	0.013
	(0.016)	(0.013)	(0.014)
ROA	0.213	0.089	0.129
	(0.224)	(0.273)	(0.242)
Leverage	-0.591**	-0.390	-0.007
	(0.239)	(0.251)	(0.311)
Log Assets	-0.177*	-0.102	-0.095
	(0.091)	(0.105)	(0.117)
Sale Growth	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.000)
Cash	0.283	-0.109	-0.015
	(0.292)	(0.358)	(0.368)
Tangibility	1.220	0.923	0.039
	(0.977)	(1.085)	(0.903)
Cap EX	-0.657	-1.560*	-0.589
	(0.713)	(0.797)	(0.912)
Log Age	0.333	0.239	0.138
	(0.401)	(0.437)	(0.336)
Customer R&D	0.076	0.078	0.015
	(0.064)	(0.049)	(0.036)
Customer Log Assets	-0.035	-0.143	-0.227**
	(0.118)	(0.111)	(0.113)
Constant	1.109	1.080	1.947
	(1.065)	(1.308)	(1.212)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,153	2,880	2,621
R-squared	0.876	0.870	0.877

Table 6: Falsification tests with artifactual assigned matched customers

This table reports the placebo regression results of the model $Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$, with matched customers. The matched customer is in the same three-digit industry as the true customer and it closest to the customer as measured by the size difference. The dependent variables are *Log Patents* in Panel A, *Log Cites* in Panel B, and *Log IE* in Panel C. 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.

Panel A: Number of Patents

	(1)	(2)	(3)
Dependent Variable	*Log Patents (t+1)*	*Log Patents (t+2)*	*Log Patents (t+3)*
Log Distance	0.004	0.004	-0.005
	(0.011)	(0.012)	(0.012)
Q	-0.002	0.003	0.015**
	(0.007)	(0.008)	(0.007)
R&D	0.352*	0.385	0.020
	(0.197)	(0.243)	(0.263)
ROA	-0.000	0.153	0.085
	(0.106)	(0.125)	(0.116)
Leverage	-0.179	-0.262	-0.266
	(0.201)	(0.173)	(0.169)
Log Assets	0.300***	0.250***	0.153**
	(0.053)	(0.062)	(0.068)
Sale Growth	0.000*	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Cash	0.134	0.042	-0.039
	(0.199)	(0.224)	(0.237)
Tangibility	0.267	0.050	0.029
	(0.458)	(0.472)	(0.469)
Cap EX	-0.231	-0.288	-0.527
	(0.433)	(0.427)	(0.445)
Log Age	0.412**	0.378*	0.330
	(0.188)	(0.217)	(0.225)
Matched Customer R&D	-0.893**	-0.379	0.565
	(0.433)	(0.463)	(0.455)
Matched Customer Log Assets	-0.011	-0.028	-0.009
	(0.019)	(0.023)	(0.024)
Constant	-0.208	0.411	0.797**
	(0.500)	(0.343)	(0.372)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	6,482	6,177	5,886
R-squared	0.865	0.857	0.856

Panel B: Number of Citations per Patent

	(4)	(5)	(6)
Dependent Variable	*Log Cites (t+1)*	*Log Cites (t+2)*	*Log Cites (t+3)*
Log Distance	-0.007	0.006	-0.008
	(0.012)	(0.011)	(0.016)
Q	-0.011	0.026***	0.009
	(0.015)	(0.009)	(0.013)
R&D	0.628	0.842**	0.514
	(0.430)	(0.388)	(0.547)
ROA	0.014	0.401	0.303
	(0.251)	(0.261)	(0.296)
Leverage	-0.104	0.064	0.142
	(0.200)	(0.198)	(0.185)
Log Assets	-0.023	0.042	0.030
	(0.090)	(0.093)	(0.098)
Sale Growth	-0.000	0.001	-0.000***
	(0.000)	(0.001)	(0.000)
Cash	0.209	-0.090	0.124
	(0.263)	(0.362)	(0.335)
Tangibility	-0.230	-0.600	-0.733
	(0.631)	(0.686)	(0.701)
Cap EX	-0.168	-0.456	0.482
	(0.780)	(0.725)	(0.924)
Log Age	-0.534**	-0.080	-0.102
	(0.247)	(0.240)	(0.254)
Customer R&D	-0.414	-0.064	-0.091
	(0.755)	(0.675)	(0.787)
Customer Log Assets	-0.028	-0.053	-0.040
	(0.036)	(0.044)	(0.045)
Constant	2.291***	2.068**	2.287**
	(0.500)	(0.848)	(1.022)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,493	3,189	2,976
R-squared	0.792	0.795	0.788

Panel C: Innovation Efficiency

	(7)	(8)	(9)
Dependent Variable	*Log IE(t+1)*	*Log IE (t+2)*	*Log IE(t+3)*
Log Distance	0.006	0.001	-0.007
	(0.012)	(0.012)	(0.011)
Q	-0.011	-0.011	0.004
	(0.013)	(0.016)	(0.012)
ROA	0.230	0.120	0.156
	(0.236)	(0.260)	(0.201)
Leverage	-0.528**	-0.424*	0.026
	(0.240)	(0.255)	(0.302)
Log Assets	-0.187**	-0.141	-0.162
	(0.088)	(0.101)	(0.114)
Sale Growth	0.002	0.001	0.000
	(0.003)	(0.002)	(0.000)
Cash	0.301	0.054	0.019
	(0.320)	(0.367)	(0.382)
Tangibility	1.241	0.846	-0.250
	(1.069)	(1.122)	(0.932)
Cap EX	-0.951	-1.582*	-0.675
	(0.774)	(0.853)	(0.941)
Log Age	0.204	0.264	0.138
	(0.407)	(0.405)	(0.328)
Customer R&D	-0.516	-0.373	0.051
	(0.663)	(0.752)	(0.755)
Customer Log Assets	-0.027	-0.024	0.011
	(0.037)	(0.038)	(0.035)
Constant	0.632	-0.322	-0.573
	(0.647)	(0.890)	(0.785)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,237	2,952	2,728
R-squared	0.872	0.865	0.870

Table 7: Falsification tests with randomized relocation timing

This table reports the placebo regression results of the model $Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$, with randomized timing of customer relocation. The dependent variables are *Log Patents* in Panel A, *Log Cites* in Panel B, and *Log IE* in Panel C. 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.

Panel A: Number of Patents

	(1)	(2)	(3)
Dependent Variable	*Log Patents (t+1)*	*Log Patents (t+2)*	*Log Patents (t+3)*
Log Distance	0.023	0.021	-0.008
	(0.044)	(0.038)	(0.050)
Q	0.001	0.005	0.015**
	(0.007)	(0.007)	(0.006)
R&D	0.424**	0.371*	-0.093
	(0.177)	(0.221)	(0.219)
ROA	0.067	0.134	0.027
	(0.097)	(0.115)	(0.104)
Leverage	-0.228	-0.280*	-0.259*
	(0.183)	(0.159)	(0.153)
Log Assets	0.320***	0.255***	0.173***
	(0.050)	(0.059)	(0.066)
Sale Growth	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Cash	0.137	-0.023	0.007
	(0.176)	(0.204)	(0.224)
Tangibility	0.351	0.004	0.052
	(0.401)	(0.414)	(0.418)
Cap EX	-0.383	-0.270	-0.545
	(0.377)	(0.379)	(0.391)
Log Age	0.373**	0.339	0.262
	(0.170)	(0.209)	(0.214)
Customer R&D	0.029**	0.019	0.019
	(0.013)	(0.014)	(0.012)
Customer Log Assets	-0.113*	-0.050	-0.073
	(0.060)	(0.066)	(0.074)
Constant	0.484	0.565	1.296*
	(0.703)	(0.635)	(0.740)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	6,994	6,694	6,380
R-squared	0.856	0.846	0.845

Panel B: Number of Citations per Patent

	(1)	(2)	(3)
Dependent Variable	*Log Cites (t+1)*	*Log Cites (t+2)*	*Log Cites (t+3)*
Log Distance	0.038	0.022	-0.026
	(0.032)	(0.039)	(0.030)
Q	-0.011	0.036***	0.011
	(0.014)	(0.009)	(0.013)
R&D	0.351	0.579*	0.166
	(0.371)	(0.343)	(0.510)
ROA	-0.017	0.211	0.294
	(0.243)	(0.261)	(0.279)
Leverage	-0.202	0.032	0.118
	(0.185)	(0.196)	(0.188)
Log Assets	-0.030	0.048	0.007
	(0.079)	(0.086)	(0.083)
Sale Growth	-0.000*	0.001	-0.000***
	(0.000)	(0.001)	(0.000)
Cash	0.096	-0.149	0.194
	(0.246)	(0.340)	(0.304)
Tangibility	-0.383	-1.044*	-0.309
	(0.610)	(0.612)	(0.674)
Cap EX	0.265	0.114	-0.013
	(0.680)	(0.691)	(0.772)
Log Age	-0.603***	-0.234	-0.322
	(0.220)	(0.222)	(0.224)
Customer R&D	-0.067	-0.015	0.002
	(0.084)	(0.090)	(0.038)
Customer Log Assets	0.049	0.046	0.076
	(0.087)	(0.085)	(0.084)
Constant	1.628**	1.338	1.515
	(0.756)	(1.001)	(1.164)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,723	3,389	3,128
R-squared	0.791	0.789	0.793

Panel C: Innovation Efficiency

	(1)	(2)	(3)
Dependent Variable	*log IE(t+1)*	*log IE (t+2)*	*log IE(t+3)*
Log Distance	-0.024	0.003	-0.015
	(0.036)	(0.029)	(0.044)
Q	-0.008	-0.008	0.009
	(0.013)	(0.017)	(0.012)
ROA	0.268	0.104	0.143
	(0.233)	(0.265)	(0.238)
Leverage	-0.499**	-0.322	0.079
	(0.244)	(0.257)	(0.308)
Log Assets	-0.181**	-0.115	-0.121
	(0.087)	(0.102)	(0.112)
Sale Growth	0.002	0.001	0.000
	(0.002)	(0.001)	(0.000)
Cash	0.284	0.034	0.101
	(0.277)	(0.350)	(0.358)
Tangibility	1.028	0.733	-0.105
	(0.974)	(1.060)	(0.907)
Cap EX	-0.709	-1.555**	-0.454
	(0.707)	(0.779)	(0.869)
Log Age	0.218	0.164	0.109
	(0.388)	(0.415)	(0.322)
Customer R&D	0.093	0.100**	0.020
	(0.066)	(0.050)	(0.035)
Customer Log Assets	-0.073	-0.141	-0.227**
	(0.115)	(0.108)	(0.108)
Constant	1.168	0.513	1.134
	(1.023)	(1.281)	(1.221)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,459	3,145	2,869
R-squared	0.865	0.862	0.871

Table 8: The mechanisms-the effects of customer R&D expense and patents

This table reports the placebo regression results of the model $Innovation_{i\tau} = \alpha + \beta_1 LogDistance_{ijt} + \beta_2 * LogDistance * LogCustomerPatent + \beta_3 * LogDistance * CustomerR\&D + \gamma_1 X_{it} + \gamma_2 Y_{jt} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variables are *Log Patents* in Panel A, *Log Cites* in Panel B, and *Log IE* in Panel C. Two interaction terms between *Log Distance* and *Log Customer Patent*, *Customer R&D* are included in the regressions. 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.

Panel A: Number of Patents

	(1)	(2)	(3)
Dependent Variable	*Log Patents (t+1)*	*Log Patents (t+2)*	*Log Patents (t+3)*
Log Distance	-0.001	0.007	0.040
	(0.055)	(0.059)	(0.063)
*Log Distance*Log Customer Patent*	-0.012*	-0.010	-0.014*
	(0.007)	(0.008)	(0.008)
*Log Distance*Customer R&D*	-0.106	-0.059	-0.130
	(0.160)	(0.226)	(0.233)
Log_C_Patent	0.143***	0.116**	0.101*
	(0.047)	(0.052)	(0.054)
Q	0.001	0.005	0.015***
	(0.005)	(0.005)	(0.005)
R&D	0.452***	0.388**	-0.073
	(0.146)	(0.163)	(0.166)
ROA	0.057	0.125	0.013
	(0.075)	(0.082)	(0.083)
Leverage	-0.214***	-0.277***	-0.254***
	(0.073)	(0.079)	(0.080)
Log Assets	0.301***	0.238***	0.161***
	(0.024)	(0.026)	(0.026)
Sale Growth	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Cash	0.125	-0.032	0.011
	(0.099)	(0.107)	(0.112)
Tangibility	0.344*	0.005	0.042
	(0.185)	(0.197)	(0.201)
Cap EX	-0.358	-0.259	-0.511**
	(0.244)	(0.257)	(0.260)
Log Age	0.364***	0.331***	0.253***
	(0.065)	(0.070)	(0.072)
Customer R&D	0.437	-0.051	0.887
	(0.826)	(1.312)	(1.349)
Customer Log Assets	-0.131***	-0.079*	-0.083*
	(0.037)	(0.041)	(0.044)
Constant	0.577	0.740	1.063**
	(0.469)	(0.504)	(0.520)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	7,000	6,700	6,386
R-squared	0.856	0.847	0.845

Panel B: Number of Citations per Patent

	(1)	(2)	(3)
Dependent Variable	*Log Cites (t+1)*	*Log Cites (t+2)*	*Log Cites (t+3)*
Log Distance	-0.060	-0.030	-0.179**
	(0.080)	(0.081)	(0.091)
*Log Distance*Log Customer Patent*	-0.011	-0.003	-0.007
	(0.010)	(0.011)	(0.011)
*Log Distance*Customer R&D*	-0.467	-0.402	-0.494
	(0.299)	(0.352)	(0.365)
Log_C_Patent	0.058	-0.049	-0.053
	(0.068)	(0.074)	(0.077)
Q	-0.012	0.037***	0.013
	(0.008)	(0.009)	(0.008)
R&D	0.331	0.548*	0.150
	(0.251)	(0.287)	(0.331)
ROA	-0.023	0.198	0.275
	(0.160)	(0.182)	(0.178)
Leverage	-0.205	0.026	0.113
	(0.126)	(0.136)	(0.137)
Log Assets	-0.013	0.053	0.010
	(0.040)	(0.043)	(0.045)
Sale Growth	-0.000	0.001	-0.001*
	(0.000)	(0.002)	(0.000)
Cash	0.113	-0.124	0.232
	(0.159)	(0.175)	(0.182)
Tangibility	-0.392	-1.008***	-0.211
	(0.328)	(0.356)	(0.373)
Cap EX	0.279	0.097	-0.129
	(0.444)	(0.475)	(0.490)
Log Age	-0.598***	-0.225*	-0.300**
	(0.115)	(0.124)	(0.123)
Customer R&D	2.477	2.041	2.284
	(1.721)	(1.882)	(1.962)
Customer Log Assets	0.041	0.063	0.102
	(0.063)	(0.067)	(0.069)
Constant	2.175**	1.733**	2.556***
	(0.851)	(0.756)	(0.789)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,725	3,392	3,131
R-squared	0.791	0.790	0.795

Panel C: Innovation Efficiency

	(1)	(2)	(3)
Dependent Variable	*log IE(t+1)*	*log IE (t+2)*	*log IE(t+3)*
Log Distance	0.075	0.084	0.083
	(0.084)	(0.084)	(0.092)
*Log Distance*Log Customer Patent*	-0.018*	-0.024**	-0.033***
	(0.011)	(0.011)	(0.012)
*Log Distance*Customer R&D*	-1.267***	-1.030***	-1.123***
	(0.322)	(0.371)	(0.376)
Log Customer Patent	0.195***	0.208***	0.242***
	(0.072)	(0.078)	(0.079)
Q	-0.008	-0.005	0.011
	(0.009)	(0.009)	(0.009)
ROA	0.259*	0.056	0.097
	(0.151)	(0.170)	(0.151)
Leverage	-0.476***	-0.303**	0.106
	(0.140)	(0.149)	(0.149)
Log Assets	-0.227***	-0.172***	-0.147***
	(0.043)	(0.045)	(0.046)
Sale Growth	0.002	0.001	0.000
	(0.002)	(0.002)	(0.000)
Cash	0.242	0.002	0.131
	(0.169)	(0.183)	(0.187)
Tangibility	0.873**	0.530	-0.295
	(0.357)	(0.385)	(0.399)
Cap EX	-0.555	-1.310**	-0.210
	(0.488)	(0.515)	(0.522)
Log Age	0.200	0.144	0.095
	(0.123)	(0.131)	(0.126)
Customer R&D	6.576***	3.880**	5.380***
	(1.844)	(1.972)	(2.002)
Customer Log Assets	-0.067	-0.181**	-0.271***
	(0.068)	(0.071)	(0.071)
Constant	0.590	0.566	0.923
	(0.900)	(0.802)	(0.805)
Year Fixed Effects	Yes	Yes	Yes
Pair Fixed Effects	Yes	Yes	Yes
Observations	3,479	3,168	2,893
R-squared	0.866	0.862	0.873

Table 9: The mechanism-the number of times a supplier's patent cites its customer's patent
This table reports the regression results of the model $Innovation_{i\tau} = \alpha + \beta LogDistance_{ijt} + \gamma' X_{it} + Year_t + Pair_{ij} + \varepsilon_{ijt}$. The dependent variable is *Log Cross Citation*, which is defined as the number of times a supplier's patent cites its customer's patent. Definitions of other 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.

	(1)	(2)	(3)
Dependent Variable	*Log Cross Citation (t+1)*	*Log Cross Citation (t+2)*	*Log Cross Citation (t+3)*
Log Distance	-0.027***	-0.027**	-0.031*
	(0.009)	(0.011)	(0.017)
Q	0.011	0.001	-0.016**
	(0.008)	(0.006)	(0.007)
R&D	-0.298**	-0.168	-0.492*
	(0.119)	(0.203)	(0.256)
ROA	-0.022	-0.019	-0.070
	(0.059)	(0.082)	(0.128)
Leverage	-0.107*	-0.138**	-0.023
	(0.056)	(0.069)	(0.073)
Log Assets	-0.049**	-0.034	-0.048
	(0.023)	(0.029)	(0.037)
Sale Growth	-0.000	0.009	-0.003
	(0.000)	(0.023)	(0.030)
Cash	-0.166**	-0.105	-0.154
	(0.078)	(0.106)	(0.150)
Tangibility	-0.043	-0.012	0.049
	(0.109)	(0.137)	(0.173)
Cap EX	0.184	0.013	0.443
	(0.250)	(0.222)	(0.288)
Log Age	0.059	0.050	-0.085
	(0.057)	(0.084)	(0.122)
Customer R&D	0.406	0.450	1.903*
	(0.337)	(0.828)	(1.115)
Customer Log Assets	0.093**	0.124**	0.197***
	(0.037)	(0.051)	(0.067)
Constant	0.040	-1.297**	-1.676**
	(0.330)	(0.584)	(0.779)
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