



Local R&D Strategies and Multi-location Firms: The Role of Internal Linkages

Juan Alcácer
Minyuan Zhao

Working Paper

10-064

Copyright © 2010 by Juan Alcácer and Minyuan Zhao

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Local R&D Strategies and Multi-location Firms: The Role of Internal Linkages

Juan Alcácer

Harvard Business School
Morgan Hall 227
Soldiers Field
Boston, MA 02163
(617) 495-6338
jalcacer@hbs.edu

Minyuan Zhao

Ross School of Business
University of Michigan
701 Tappan St., E4616
Ann Arbor, MI 48109
(734) 647-6978
myzhao@umich.edu

First version: January 2006

This version: March 2009

Local R&D Strategies of Multi-location Firms: The Role of Internal Linkages

ABSTRACT

This study looks at the role of firms' internal linkages in highly competitive technology clusters, where much of the world's R&D takes place. The leading players in these clusters are multi-location firms that organize and integrate knowledge across sites worldwide. Strong internal links across locations allow these firms to leverage knowledge for competitive advantage without risking critical knowledge outflow to competitors. We examine whether multi-location firms increase internal ties when they face appropriability risks from direct competitors. Our empirical analysis of the global semiconductor industry shows that when leading firms co-locate with direct market competitors, innovations tend to be quickly internalized, and are more likely to involve collaboration across locations, particularly with inventors from the firm's primary R & D site. Our results suggest that R&D dynamics in clusters are heavily influenced by multi-location firms with innovative links across locations, and that future research on technology innovation in clusters should account for these links.

KEYWORDS: technology clusters, knowledge spillover, internalization, appropriability.

1. Introduction

Firms co-locate in technology clusters for many reasons. These clusters -- geographic concentrations of firms and other institutions engaging in closely related R&D activities -- might attract firms using the unique advantages of a location, such as proximity to research universities, favorable government policies, and an abundance of human capital. In other cases, clustered firms emerge as spinoffs of existing companies in the cluster. In fact, most innovation in technology-driven industries occurs in technology clusters.

Meanwhile, clusters are often characterized by “vigorous competition among locally-based rivals” (Porter 1998), which could compromise a firm's competitive advantage. Because much knowledge is tacit, effective knowledge transfer often requires frequent interpersonal interactions, which are more likely to happen with geographic proximity (Jaffe *et al.* 1993; Audretsch and Feldman 1996). Technological clusters also facilitate labor mobility (Almeida and Kogut 1999), an additional mechanism of knowledge flow across organizations. Firms and the R& D community at large benefit from the knowledge inflows that mobility and informal interactions bring, but unintended knowledge outflows to competitors can erode the competitive edge of industry leaders.

What enables firms to benefit from location-specific advantages without endangering their technological edge? We address this question by focusing on an important factor in cluster dynamics: the leading players in technology clusters are often geographically dispersed organizations with R&D at multiple locations. Large multi-location firms are known for their ability to mobilize and integrate knowledge on a global basis (Bartlett and Ghoshal 1990). Thus, to understand R&D dynamics in a cluster, we must recognize that a firm located in a particular cluster may also be part of an extended network, with its operation strategically integrated across multiple locations and multiple business lines. The innovation strategy of IBM in Cambridge, Massachusetts, for example, is intricately linked with the company's eight other R&D labs and hundreds of facilities worldwide.

Previous studies show that internal linkages across a firm's geographically dispersed units can improve knowledge absorption and integration (Gupta and Govindarajan 2000). Strong linkages across a

firm, as evidenced by the collaboration of inventors across distance, increase the absorption of external knowledge at dispersed locations (Lahiri 2003), facilitate the transfer of local knowledge back to the parent firm (Zhou and Frost 2002), and improve the overall quality of innovation (Singh 2007).

Less attention has been paid to the role of internal linkages in the appropriability of knowledge. Since specialized and co-specialized complementary assets are critical to the commercialization of an innovation (Teece 1986), firms can minimize their losses from knowledge outflow to competitors by strategically increasing the interdependence among their global network of subsidiaries (Feinberg and Gupta 2009). Local imitators without a similar integration mechanism would be unable to make use of any knowledge gained. Studying multinational firms' R&D activities in countries with weak intellectual property rights (IPR) protection, Zhao (2006) argues that a firm can utilize its internal organization to substitute for the weak external institutions. When a firm's R&D network spans multiple locations, at each location it can develop technologies that closely relate to the firm's internal resources residing elsewhere around the world. As long as the innovating firm can integrate the locally developed knowledge more efficiently than potential imitators, it can take advantage of the low cost in weak IPR countries and still gain from innovation. This is in line with the disaggregation strategy suggested by Liebeskind (1996) as a mechanism to protect against knowledge outflows. Sanna-Randaccio and Veugelers (2002) develop a theoretical model suggesting that firms would locate in clusters where competitors are present only if they have internal organizational structures to tightly control the innovations generated there. A well-integrated organization with strong internal linkages may also detect and absorb internal innovations more efficiently, thus increasing their lead-time over competitors -- another important mechanism of appropriability (Levin *et al.* 1987; Cohen *et al.* 2000).

The multi-dimensional relationships among local entities (Cohen 1995, 230) allow us to separate appropriability incentives from learning incentives among multi-location firms. Firms in a technological cluster may share similar technological backgrounds or even engage in patent races, but they do not necessarily compete in the same product market. Industry-specific market information and other complementary resources reduce the risks associated with knowledge exchanges, allowing symbiotic

relationships to develop. If internal linkages are purely mechanisms of knowledge integration, we should observe more internal linkages in regions with greater learning opportunities, e.g., with more neighboring firms in the same technological fields. If internal linkages serve as a hedge against knowledge outflows, they should be used more extensively when neighboring firms share the product market.

Examining the global semiconductor industry from 1998 to 2001, we find supportive evidence for both the learning mechanism and the appropriability mechanism of internal linkages, with much stronger results for the latter. Specifically, when surrounded by direct competitors, the technology leaders in a cluster favor technologies that can be quickly developed internally, and more of their R&D projects involve researchers from other locations, particularly from primary R&D sites. With comprehensive analysis of the competitive environment, this study extends the findings of Zhao (2006) and suggests that internal linkages across a firm protect firm knowledge from appropriation not only in weak IPR countries, but also in risky competitive environments in general. In other words, internal linkages are one of the dynamic capabilities that enable firms to embrace opportunities in technology clusters (Teece *et al.* 1997).

The rest of the paper is organized as follows. Literature review and theory development are in the next section, followed by the description of data sources and empirical design in Section 3. The empirical results and robustness tests are presented in Section 4. Section 5 concludes.

2. Theoretical Analysis

2.1 Technology Clusters and Multi-Location Firms

According to Porter (1998), clusters are a prominent feature in the landscape of every advanced economy. Starting with seminal work by Marshall (1920), researches have shown that firms in an industry cluster benefit from knowledge outflows to competitors, access to specialized labor, and access to specialized intermediate inputs. Among the various activities along the value chain, R&D activities benefit the most from knowledge transfer between competitors, and thus show the highest level of concentration (Audretsch and Feldman 1996; Alcácer 2006). Geographic proximity enables frequent interpersonal interactions through existing social networks (Almeida and Kogut 1999) and local

institutions (Gilson 1999; Stuart and Sorenson 2003), which facilitate the transfer of tacit knowledge in technology clusters.

Knowledge, however, flows in both directions. Knowledge flowing into the firm (knowledge inflow) is likely to make R&D investments more productive, and thus raise the incentives to invest in it, but knowledge flowing out of the firm (knowledge outflow) may hinder the firm's ability to appropriate value from its own innovations, thus lowering its incentive to conduct R&D in clusters (Furman *et al.* 2006). In particular, losing knowledge to nearby competitors erodes the competitive advantage held by leading firms. Firms can move away from clusters to protect their cutting-edge technologies (Shaver and Flyer 2000), but this option may not be sustainable or desirable for two reasons. First, even if a leading firm decides to locate apart, it has little control over the subsequent location decisions of competitors or the emergence of new firms. To the extent that other firms have incentive to cluster around industry leaders, geographic distance offers only temporary protection against knowledge outflow. Second, there may be crucial resources in the cluster that the firm relies on, such as the talent pool from a local university. Relocation would seriously compromise the firm's long-term competitiveness in the industry. Hence, protecting innovations from nearby competitors is a strategic consideration leading firms can't avoid.

One feature that industry leaders can take advantage of is their geographically dispersed, yet closely integrated, innovation networks. The literature of technology clusters traditionally treats all local entities as stand-alone organizations. As a result, interactions among local competitors have been examined without much consideration of firms' extended organization. At the same time, most of the leading firms in high-tech industries are large firms with R&D activities in multiple locations – if not countries. As emphasized by Pisano (2006), an industry's methods of appropriation are created by the strategic decisions of firms in that industry. Hence, the strategic allocation and integration of R&D activities by multi-location firms will have important implications for firms' interactions in technology clusters.

2.2 Internal Linkages and Knowledge Appropriation

Researchers have long recognized firms' internal linkages as effective mechanisms of knowledge absorption and integration. By establishing interactions across divisions, intra-firm ties facilitate the accumulation and integration of knowledge (Bartlett and Ghoshal 1990; Kogut and Zander 1993). Empirical evidence shows that strong internal linkages – evidenced by collaboration among inventors across distances – are conducive to the absorption of external knowledge (Lahiri 2003) and the knowledge flow from foreign subsidiaries to the parent companies (Zhou and Frost 2002). Furthermore, such linkages also affect innovation quality. Singh (2007) shows that geographic dispersion of R&D, once accompanied by sufficient cross-regional ties among researchers from different R&D units, is associated with an improvement in innovation quality.

In this study, we argue that internal linkages also help firms extract maximum value from their innovations by facilitating learning and knowledge integration. Local innovations can be quickly leveraged by other parts of the firm, leading to stronger competitive position in the product market. Knowledge assets are cumulative in nature (Dierickx and Cool 1989), and most products are the result of a long sequence of technological improvements (Vickers 1986). "Because the knowledge embedded in technological innovations is often tacit, the potential value, the potential value of an innovation is not always straightforward to outsiders (Arora *et al.* 2001). The innovating firm, with its hands-on experience in the R&D process and comprehensive understanding of the context, is able to develop its innovation without first completing a costly learning process (Mansfield *et.al.* 1981). If a firm can build on its new technologies more efficiently than potential imitators, it may gain crucial lead-time for value appropriation (Cohen *et al.* 2000).

Second, strong internal linkages also lead to higher interdependence between the local unit and firms' organizational expertise, creating knowledge that is hard to replicate by local competitors. Because specialized and co-specialized complementary assets are critical to the successful commercialization of an innovation (Teece 1986; Anand and Galetovic 2004), firms can minimize the risk of knowledge outflows by strategically increasing internal linkages. On the one hand, the firm-specific nature of corporate R&D

leads to heterogeneity in R&D applications and increases the appropriability of R&D returns (Helfat 1994). On the other hand, technologies generated by local subsidiaries are less attractive to neighboring firms if these competitors don't also have access to complementary information residing outside the cluster (Zhao and Islam 2006). Rajan and Zingales (2001) also show the importance of interdependence in knowledge appropriation. Using a theoretical model, they explain why flat hierarchies – in which all division managers are required to collaborate with a central unit at the top – are ubiquitous in human capital-intensive industries such as legal and consulting services. Because of the intangible nature of firm resources, property rights protection is difficult to enforce. Yet, if the firm can increase everyone's dependence on the center office by controlling access to certain key resources, the risk of expropriation is greatly reduced. Similarly, Liebeskind (1996) points out that disaggregating tasks gives firms an advantage in knowledge protection, especially when reinforced by spatial isolation. In other words, the risk of knowledge outflows is reduced if the divisions of a firm are highly dependent upon each other.

Based on the above discussion, we argue that firms will appropriate the most value from R&D in technology clusters when they internalize their innovations better and faster than nearby competitors – thereby reducing the damage of knowledge outflow and gaining important lead time over competitors. Moreover, the importance of such strategies increases with the intensity of local competition. In the next section, we will explicitly test the application of two internalization mechanisms in highly competitive technology clusters.

3. Empirical Design

3.1 Sample

Our empirical setting is the worldwide semiconductor industry from 1998 to 2001. We choose this industry for several reasons. First, innovation is a key factor for success in semiconductors. Firms invest relentlessly in R&D to introduce new products and improve production processes (Stuart 2000). Moreover, semiconductor firms routinely patent their innovations, and patent data have been used to trace the traits and geographic distribution of innovation. Second, the benefit of knowledge transfer between

firms has been shown to drive agglomeration in the industry (Saxenian 1994; Fleming *et al.* 2006). High levels of geographic concentration also suggest that semiconductor firms have already developed strategies to manage knowledge outflows. Third, this is a truly global industry: leading firms operate at multiple locations around the world, and there are significant differences between firms in terms of product markets, R&D portfolios, positions in the value chain, and geographic locations. Firms range from industry giants that participate in activities throughout the value chain to small enterprises that specialize in design (known as fabless) or testing, and from large multinational firms to small local firms. Other players, such as universities, national laboratories, and firms from other industries (e.g., aerospace and chemicals) also conduct active R&D in semiconductors. Such heterogeneity allows us to identify the effect of different competitive environments on firms' appropriability strategies and allocation of R&D projects.

We build our dataset from four sources. First, we identify innovating firms using patent data from the *Derwent World Patent Index* (DWPI), a well-recognized dataset that encompasses more than 30 million patent documents from 41 patent-issuing authorities worldwide, and we rely on *Derwent's* technological classification¹ to obtain the universe of semiconductor patents. Patent data include innovations that occur outside of the R&D facilities, thus are more inclusive than the number of labs or the amount of R&D spending. Information from semiconductor patents applied between 1998 and 2001, and granted between 2001 and 2004, results in a sample of 60,880 patents.

Many of these patents are linked to the same innovation, with exactly the same inventors, assignees and abstracts. Multiple patents per innovation can occur either because patents are filed in multiple countries or because an application in a given country spins out multiple patents. For example, 16% of patents granted by the U.S. Patent and Trademark Office (USPTO) in our sample are duplicates. Thus, we

¹ DWPI applies a consistent classification system to all patents. Classes used in this study are U11 (semiconductor materials and processes), U12 (discrete devices), U13 (integrated circuits) and U14 (memories, film and hybrid circuits). For more details, see <http://scientific.thomson.com/support/patents/dwpi/ref/reftools/classification>.

follow Gittelman and Kogut (2003) and use families of patents as our unit of analysis. Each family encompasses patents granted in all countries that are identical in terms of technology, inventors, and locations, and differ only in the scope of their claims. The final sample consists of 23,675 patent families² whose assignees are American and foreign firms, universities, and government- and industry-sponsored research labs. For the 300 patent families with more than one assignee, all assignees (and not only the first one) are considered. Patents granted in the U.S. represent 46% of all the patent family members, followed by those in Europe (17%), Japan and Korea (7% each), and Taiwan (6%).

We supplement this initial sample with directories of semiconductor plants, fabless companies, and institutions behind scientific publications. Information about plants comes from the quarterly datasets of the *World Fab Watch* provided by the Strategic Marketing Association, from 1998 to 2001. The datasets encompass manufacturing facilities for a wide range of products: memories, microprocessors, generic and specific chips, etc. Information on fabless companies is obtained from the Gartner Group's annual *Directory of Fabless Semiconductor Companies* for the same period. To assess the scientific activities in the local community, we extract from *ISI Web of Knowledge* all journal publications in the sample period that use "semiconductor" or "semiconductors" as part of their keywords. These four data sources provide a comprehensive map of the industry at multiple levels: innovation (23,675 patent families), production (974 plants), research (26,581 scientific publications), and development (549 fabless companies).

Because we treat every multi-unit firm as an integral entity, and because internal organization is a central concept of this study, we put extra effort into identifying the ultimate parent for every entity in our sample. First, for each year, we match the patent assignees, plants and fabless companies to firms in the corresponding *Directory of Corporate Affiliations* (DCA), an annual database that records corporate ownership for more than 200,000 private and public firms worldwide. Second, for organizations not identified in DCA, we search the *Dun and Bradstreet Million Dollar Database* to obtain affiliation information. Finally, we check affiliation changes through SDC Platinum, company websites, and various

² Besides patents, these families also include 29,491 patent-related documents such as PCTs.

industry publications. The above steps map the 4,125 assignees in the sample to 2,217 unique organizations. Fabless firms and manufacturing firms that do not own patents add 721 additional organizations to our sample.

While we use data for all organizations to characterize local environments, our analysis of R&D strategies focuses on 16 innovating firms, or the top 1% of the industry in terms of patent output³. The reason for focusing on these firms is that most of the semiconductor industry has the typical features of an oligopoly industry, where the top 1% of firms represent 50% of the patent output and 40% of the plants operating in this period. With the cost of developing new chips and building new manufacturing plants running into the billions, there is a clear divide between industry leaders and everyone else – and the gap is getting larger, according to IC Insights. Therefore, semiconductor industry leaders should have qualitatively different innovation strategies than the thousands of industry followers. The composition of the sample is similar to those in previous studies of the semiconductor industry (Stuart and Podolny 1996; Henisz and Macher 2004; Ziedonis 2004). As part of our robustness checks, we replicate our analyses using two alternative samples – composed of the top 5% and 10% of firms in terms of patent families – and obtain similar results.

3.2 Cluster Definition

Defining technology clusters is a crucial element of our empirical setup. Instead of relying upon predetermined administrative boundaries, such as states or metropolitan areas, we apply a mathematical algorithm that uses latitude and longitude data to identify technological clusters. We do this for two reasons. First, there is no single administrative unit defined across all countries. We have to either focus on a specific country (e.g., the U.S.), which fails to capture important features of global firms, or use a mix of different geographic units (e.g., states in the U.S., prefectures in Japan, and provinces in Europe), which may create unexpected country biases. Second, technological clusters do not necessarily follow

³ The 16 firms are AMD, Intel, IBM, Texas Instruments, Hitachi, Matsushita, NEC, Siemens (including Infineon), Toshiba, Mitsubishi, Samsung, Micron, Fujitsu, TSMC, Hyundai, and STMicroelectronics.

predetermined administrative boundaries, which is clear after a quick inspection of inventor locations in, for example, the northeastern U.S. or central Japan. One administrative unit may encompass multiple clusters, while one technological cluster may expand across several administrative lines.

In this study, we define clusters by the actual distribution of inventor locations, following a three-step approach. First, we identify the location of each element in the sample (i.e., a patent inventor, plant, fables company, or scientific publication), and match the locations to two comprehensive sources of geographic names. For U.S. locations, we obtain latitude and longitude information for all 38,261 locations in the country from the *Geographic Names Information System* (GNIS) of the U.S. Geological Survey. For foreign locations, we use the *Geonet Names Server* (GNS) of the National Geospatial Intelligence Agency. Besides its wide coverage of 5.5 million location names worldwide, the GNS dataset uses phonetic variations to capture spellings from a different alphabet (as in Asian countries) and from an alphabet with extra characters (as in Scandinavian and Slavic countries). Ambiguous matches are checked manually by native residents from various countries and areas. As a result, we are able to assign latitudes and longitudes to 38,926 out of the 38,952 foreign locations in the original sample.

In the second step, we develop a mathematical algorithm to identify geographic clusters using the latitude and longitude information. Clusters are defined not only by the geographic distance among locations – as in many other traditional clustering methods – but also by the variations in inventor density in neighboring areas. For example, a rapid decrease in density may signal the end of a cluster, and a continuous level of inventor density may signal a long or irregularly shaped cluster. Accordingly, the algorithm assigns two locations to the same cluster if there is a continuity of high-density locations between them, despite their geographic distance. In contrast, two locations separated by a stretch of low-density areas may be identified as two distinct clusters, even if they aren't far apart. Our clustering algorithm offers the additional advantage of having the number of clusters emerge naturally from the data, instead of being set arbitrarily *ex ante*. This method produces 304 geographic units.

Finally, plants, fables companies, and publications are assigned to the geographic units defined from the patent data. In most cases, they fall within an existing geographic unit. For each location that

falls out of all existing units, we calculate its shortest distance to them. The location is considered part of the closest cluster if the minimum distance is less than 15 miles⁴. Otherwise, the unassigned locations are again clustered with the same algorithms as we use for the patent locations. For the main sample, 6 and 28 geographic units were added by fables and plant data, respectively⁵.

3.3 Dependent Variables

3.3.2 Internalized Value

A key concept in this study is the extent to which an innovation creates value for the innovating firm. While there is no direct measure of value, technologies highly dependent on internal resources are more likely to be utilized and further developed within the firm. Trajtenberg *et al.* (1997) propose self-citations, defined as “the percentage of citing patents issued to the same assignee as that of the originating patent,” to measure the “fraction of the benefits captured by the original inventor.” Hall *et al.* (2005) also suggest that citations to patents belonging to the same firm represent internalized knowledge transfers, bolstering the firm’s competitive advantage. Hence, we use forward self-citations as a proxy for the value new technologies bring a firm. Specifically, we define the variable *self_citation_p* as the number of self-citations among all citations received by patent family *p*; citations to a patent family sums citations by all its members. Because we are interested in firms as integrated organizations, any citations among affiliated organizations are considered self-citations.

A common critique of citation-based measurements is the unknown nature and extent of citations imposed by patent examiners (Jaffe *et al.* 2000). Recent research reveals that examiner citations account for 66% of all citations in an average patent, which may bias empirical tests (Alcacer and Gittelman 2006; Sampat 2009). To avoid this problem, our main models are estimated using citations listed by inventors

⁴ We also tried 20, 25, and 30 miles, with very similar outcomes.

⁵ Note that the geographic units identified are not necessarily technology clusters, which are units with high innovation densities. For convenience, we use “cluster” and “geographic unit” interchangeably whenever there is no concern of confusion. The analysis is replicated with a hierarchical clustering algorithm in robustness checks.

only. In our sample, about 30% of the patent families that receive at least one inventor citation also have at least one self-citation. The number is 38% when both inventor and examiner citations are considered. For robustness checks, we also repeat our analysis using all citations to a patent regardless of their source.

Note that the citation measures capture both the intensity and speed of citations. As our observation window ends in December 2006, any citations that occur after that date are not included in the sample. Hence, if we observe a large number of citations on a patent, it is either because the patent is widely cited in general or because the citations occur fast enough to be captured in the short observation window. In either case, a high self-citation ratio suggests that the innovating firm internalized the focal technology.

Geographically dispersed R&D in a multi-location firm makes it more difficult for local competitors to access the technology know-how residing in the firm's other subsidiaries, thus reducing knowledge outflow (Sanna-Randaccio and Veugelers 2002). Assigning R&D projects to teams spanning multiple clusters can create links within the organization that not only enhance appropriability, but also facilitate the transfer of local know-how throughout the organization (Lahiri 2003). Thus, we define *cross_cluster_{ict}* as the number of patent families per firm-cluster-year for which the inventors are from at least two different clusters.

Furthermore, we differentiate the firms' primary R&D sites – locations where firms conduct most of their R&D – from their peripheral R&D facilities. We then identify cross-cluster links that involve only the primary R & D site, and those connecting peripheral locations. As shown in Table 1, the average number of cross-cluster links per firm-cluster-year is 4.4. Cross-cluster teams involving primary R & D sites are more common than teams linking peripheral locations (1.65 vs. 1.38).

3.4 Independent Variables

We follow two dimensions – technology space and product market – to characterize the competitive environment at the cluster-year level. Along the technology space, competitors are defined generically as organizations that innovate in the semiconductor field. The variable *innovators* represents the number of unique assignees with semiconductor patents in a given cluster-year. We then classify assignees into two groups: *innovators_profit* and *innovators_nonprofit* to capture the number of for-profit and nonprofit

assignees, respectively. In addition, we use the status information on patent applications⁶ to further classify for-profit assignees into small or large entities, thus creating the variables *small_innovators* and *large_innovators*. In the case of nonprofits assignees, we manually classify them into three groups: universities (*universities*), government agencies (*govt_innovators*), and other nonprofits such as research centers sponsored by industry associations (*other_nonprofit*).

Along the second dimension, competitors are defined as firms that share the same product-market. For every focal firm in our sample, we rely on *Hoover's Online* to identify its industry (four-digit SICs), market segments within semiconductors⁷, and the names of direct competitors. Then we count the number of for-profit assignees in the same industry (*in_industry* and *not_in_industry*), in the same market segment (*in_segment* and *not_in_segment*), or on the list of direct competitors (*competitors* and *not_competitors*). The self-reported competition data from Hoover's serve our purpose well, since managers make strategic moves based upon perceived competition in a technology cluster.

We complete the characterization of local innovation environments with three more variables: *plants_in_cluster*, *fables_in_cluster*, and *publications_in_cluster*, which represent the numbers of plants, fables companies and publications per cluster-year. In addition, we use two dummy variables, *with_plant* and *with_fables*, to indicate whether a particular firm has plants or fables units in cluster *c* and year *t*. At the firm level, we include two variables, *patents_semi* and *patents_total*, to capture the number of patents that a firm has through year *t* in semiconductors and in all technological classes, respectively. Our focal firms have on average 200 semiconductor patents and 1,295 patents in all technology categories. Table 1 presents the descriptive statistics for all dependent and independent variables used in the empirical tests.

⁶ The USPTO uses industry-specific parameters such as number of employees and revenues to grant small firm status to assignees. For details see http://www.uspto.gov/web/offices/pac/mpep/documents/appxr_1_27.htm.

⁷ Hoover's reports 13 segments under semiconductors, including memory chips and modules, microprocessors, etc.

3.4 Methods

To identify firms' strategic organization of R&D projects across locations, we compare the technologies developed in different local environments, controlling for firm characteristics. Specifically, the two dependent variables – *self_citation*, and *cross_cluster* – are the two dimensions that characterize the appropriability strategies described in the previous section.

Thus, we estimate two basic equations, one for each dependent variable, in the following form:

$$DV_{ict} = C_{ict} + X_{ict} + Y_{ct} + \zeta_t + v_i + \tau_{ctry} + \varepsilon_{ict} \quad (2)$$

where C_{ict} is a vector of cluster-specific variables capturing the competitive environment faced by firm i in cluster c and year t , X_{ict} is a vector of firm-specific variables characterizing firm i in cluster c and year t , and Y_{ct} is a vector of location characteristics in year t . ζ_t and v_i are two sets of dummy variables for year and firm fixed effects, respectively. Variations in country-specific intellectual property right regimes are controlled by the country dummies τ_{ctry} , and ε_{ict} is the error term.

Note that the analysis for *self_citation* is conducted at the innovation level (i.e., patent family), while the analysis for *cross_cluster* is at the firm-cluster level. As *self_citation* and *cross_cluster* are both count variables, negative binomial models are used for the estimations⁸.

4. Empirical Results

Table 2 presents the results of estimating *self_citation*, using negative binomial models. Because the dependent variable is the number of self-citations received by the focal patent, and the exposure variable is the total number of forward citations, we are essentially examining the patent's self-citation ratio. OLS regressions with self-citation ratio as the dependent variable produce very consistent results.

The total number of innovators in the cluster does not seem to have a significant impact on internalization, even if we only consider for-profit innovators. The effect of competition starts to emerge when we distinguish large, for-profit innovators from small ones. An increase in the number of

⁸ The exposure variables are total citations and total patents, respectively.

neighboring large firms increases the self-citation ratio, while the opposite is true for neighboring small firms. The effect of local competition is more evident when firms compete in the product market than it is when firms employ similar technology in different markets. Across various specifications of market competition, the coefficient on the number of local competitors is positive and significant. The more market competitors there are in a cluster, the more likely firms are to self-cite patents they develop there. To the extent that self-citations proxy for internalized value, this finding supports our argument that in highly competitive environments, firms are more likely to share technology development across the firm.

Meanwhile, the presence of nonprofit innovators has little impact on the degree of internalization. Without direct market competition, these nonprofit institutions create a more open atmosphere in the local cluster. An alternative explanation for this phenomenon is that firms choose to locate in close proximity to universities or government laboratories for the purpose of seeking knowledge. Intensive internalization may negatively affect the firm's ability to absorb external information. Not surprisingly, the coefficient of *patents_semi* is positive and significant; the larger the patent pool in the technological domain, the more likely that later citations are made to that pool. The coefficient of *with_plant* is still positive and significant, indicating that technologies closely linked to manufacturing processes are more firm-specific.

Note that the high self-citation ratios in competitive clusters are not due to the low intrinsic value (small denominator) of these patents. When running the same regressions with total number of citations instead of self-citations as the dependent variable, none of the coefficients associated with competitive environments are significant. To further verify this point, we compare the number of self-citations and the number of total citations – commonly used as measure of patent quality – across various competitive environments. Specifically, we use *Hoover's* data on direct market competition to define four quartiles, with Quartile 1 indicating the clusters with the highest number of direct competitors and Quartile 4 with the lowest number of competitors. As shown in the two panels of Table 3, while there are significant differences in self-citations across quartiles – with on average 1.72 self-citations found in clusters with most competitors vs. 0.52 self-citations in the least competitive clusters – we find no statistical evidence that patent quality varies across quartiles. Together, these findings suggest that firms do change the type

of innovation performed depending on local environments. Innovation produced in clusters with a strong presence of direct competition is more tightly intertwined with the firm's internal knowledge base.

While self-citation ratio is a good proxy for knowledge appropriation by the innovating firm, the findings in Table 2 can also be interpreted as learning. Innovating firms are eager to absorb knowledge from local competitors, internalize it, and spread it across firm locations. As a result, patents developed in highly competitive clusters feature higher self-citations. To argue that the high self-citation ratios in highly competitive clusters are a product of knowledge protection rather than a product of learning, we must show that citations among competitors are reduced within the cluster. Table 4 illustrates the backward citation patterns within and across clusters by various entities. Throughout the three alternative definitions of competition, firms cite competitors relatively less within a cluster than across clusters. For example, the ratio of competitors' citations to other firm's citations is 1.47 outside the cluster versus 0.87 within the cluster when competitors are identified at the industry level; 2.5 vs. 1.62 when identified at the segment level and 2.09 vs. 1.63 when measured at the self-revealed competitor level. If the positive and significant relationship between self-citations and level of competition in a cluster was driven by learning opportunities presented by competitors, one would expect that citations to local competitors would be higher than citations to remote competitors, or at least that there would be no difference. The figures in Table 4 suggest that firms' internalization strategies in competitive clusters are more likely designed to prevent knowledge outflow to competitors.

Table 5 shows the regression results of *cross_cluster* with negative binomial models. Models (1) through (5) use the total number of local patents as the exposure variable, so we essentially test the percentage of local patents developed by cross-cluster teams. The positive coefficients of *innovator_profit*, *in_industry*, *in_segment* and *competitors* suggest that the presence of competing organizations increases the tendency to use cross-cluster teams. As with the analysis of self-citations, the number of nonprofit innovators has no effect on the use of cross-cluster teams.

Most of the results with control variables follow the same pattern as in the previous tables. The presence of a plant in the cluster increases the use of cross-cluster teams, probably a reflection of

production-related projects that require inputs from local engineers and R&D personnel at the firm's primary R& D site. Note that we add a new control variable, *primary_cluster*, to indicate whether the cluster is the primary R&D site for the firm. One would expect a centralized system to exert more control over geographically dispersed innovation by tightly connecting them with a primary R&D center. As expected, the coefficient of *primary_cluster* is positive and highly significant.

We further explore cross-cluster links between primary and peripheral locations. Models (6) through (10) show the results with a new dependent variable, *to_cluster_{ict}*, which counts the number of patents in the non-primary cluster *c* with at least one inventor located in firm *i*'s primary cluster. The exposure variable used for this estimation is the number of firm *i*'s cross-cluster patents in cluster *c* (*cross_cluster_{ict}*); hence, we essentially explore the percentage of cross-cluster patents linked to the primary cluster. The results are very consistent with previous findings. That is, an increase in the number of competitors in the peripheral clusters increases the percentage of cross-cluster links connecting to the primary cluster, and this effect is stronger when the competitive environment is measured by product-market competition. Therefore, we find that local competition increases cross-cluster links, and that the increase is due, specifically, to connections with the firms' primary R & D site.

4.4 Robustness Checks

The above findings are consistent with our hypothesis that R&D projects in competitive clusters are developed more internally, and are more likely to involve teams spanning multiple locations. Next, we conduct a series of robustness tests using alternate samples, variable definitions, and estimation techniques.

First, we re-estimate all models with a different method to define clusters: hierarchical clustering with centroid linkages. This method begins with each location as a separate group. Then two clusters with the shortest Euclidian distance are combined into one, whose new geographic coordinates are the mean longitude and latitude of all locations in the group. This process is repeated until a large hierarchical tree is generated that includes all locations. We designate the number of clusters in each region to accommodate a wide variation in local densities. This process produces 187 distinct geographic units. The

coefficients obtained with the hierarchical clustering method are similar in sign, significance and magnitude to those in the previous tables.

Second, we repeat the analysis on self-citation ratios using both inventor and examiner citations. Recent research suggests that high levels of examiner citations are associated with low quality patents (Alcácer and Gittelman 2006; Sampat 2009). Therefore, including these citations adds a new set of observations – patents whose citations are 100% examiner-imposed – that may represent inferior innovations. The results using citations from all sources are similar in magnitude and sign to, but weaker in statistical significance than, those in Table 2.

Finally, we estimate the models with cluster-fixed effects to control for unobservable factors at the cluster level. Due to the large number of dummy variables for firms, years, countries, and clusters, some models fail to converge. For those models that do converge, which constitute a significant majority, the competitive measurements based on product market, especially those related to direct competition, come up with coefficients that are statistically significant with the expected signs. Since any location-specific variations are controlled for with the cluster dummies, the results strengthen our belief that firms tailor their R&D strategies to their competitive environment.

5. Conclusion

While geographic co-location has obvious benefits for firm innovation, it can also have serious drawbacks. We explore how firms tap into technology clusters' rich resources while protecting the value of their innovations. Our empirical findings suggest that leading firms organize their R&D activities differently when facing local competitors. A multi-location firm may reduce imitation risks by allocating less vulnerable projects to clustered areas, by incorporating local innovations quickly into its global knowledge base, and by using cross-cluster teams to strengthen control over locally developed technologies. We also find that firms' strategic responses vary depending upon the characteristics of nearby organizations. We find strong evidence of strategic behavior when neighboring firms share the same product market, but not when they overlap in the technological space.

Of course, there remain several limitations in this study. For example, this is a one-industry study with 16 focal firms; further analysis with more industry contexts will make the conclusions more generalizable. In addition, the self-citation measures we use to capture internalization may also reflect firms' knowledge acquisition strategies. Since the direct market competitors of our focal firms are most likely leading innovators, the stronger internal ties we observe among focal firms may simply reflect the learning process. Although we used two approaches to rule out this alternative explanation -- using distinct measures (the presence of market competitors and technological innovators) to proxy for competition effect and learning effect respectively, and looking at within-cluster and across-cluster citations to competitors -- it is difficult to tease these out completely.

Despite these limitations, we believe our study sheds light on important aspects of location and innovation strategies. For firms making location decisions, this study shows that highly competitive technology clusters are not a forbidden land for industry leaders. Knowledge outflows to competitors could threaten a firm's technological leadership, and may even erode its competitive advantage, but these threats shrink when a multinational spreads its R&D efforts strategically across firm locations. The risk of exposing certain technologies to local competitors is also low if local competitors lack the capabilities to absorb those technologies. Taking it one step further, because the potential for knowledge outflows from industry leaders will prompt small firms to cluster nearby, avoiding technology clusters is hardly an option for the most technologically advanced. Strategic organization of R&D activities becomes crucial.

Policy makers eager to nurture local high-tech industries often use various incentives, such as tax breaks, to attract firms to conduct R&D locally. At the same time, government has little influence on how R&D is actually conducted. With local projects closely intertwined with the firms' global research agenda, the same R&D budget or R&D intensity may generate very different knowledge outflows to the local community. It would be interesting for future research to examine the features of local environments, such as labor mobility, that facilitate not only R&D investments, but also active learning across firm boundaries.

This study also points to several avenues for further inquiry. First, although the mechanisms explored in this paper are based on multi-location firms, the need to appropriate economic rents from proprietary innovation applies to any firm or organization. More research is needed to understand other appropriability mechanisms that do not rely specifically on multiple locations.

Second, the strategies discussed in this study are based on a well-established set of internal routines and organizational skills that facilitate the transfer and integration of geographically dispersed knowledge. Obviously, not every firm can achieve the strategic allocation of R&D resources with enough efficiency or cost effectiveness. Hence, it is important to understand how firm heterogeneity affects the applicability of these strategies, and how various internal organizational structures influence firms' abilities to absorb, transfer and protect knowledge in technology clusters.

Third, our arguments revolve predominantly around competition and have excluded the possibility of inter-organizational cooperation. However, there are frequent project collaborations, strategic alliances, and industrial associations among semiconductor firms, universities, and other research institutions. Cooperative arrangements are even observed between direct market competitors. Such arrangements may affect the nature of R&D in a location and the appropriability mechanism at play.

Finally, in the semiconductor industry, as in many other high-tech industries, R&D is fragmented across the value chain and, in some cases, outsourced to specialized firms (Arora *et al.* 2001). In such circumstances, knowledge flow across organizational boundaries is necessary and desirable. Moreover, firms' abilities to allocate resources and exercise strategic internalization are limited once innovation moves outside the firm. Therefore, we need to better understand how firms protect and extract value from innovations developed within permeable, changing, and diffuse firm boundaries.

References

Alcácer, J. 2006. Location choices across the value chain: How activity and capability influence co-location. *Management Sci.* **52**(10) 1457–1471.

- Alcácer, J., M. Gittelman. 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *Rev. of Econom. and Statist.* Forthcoming.
- Almeida, P., B. Kogut. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Sci.* **45** 905–917.
- Anand, B.N., A. Galetovic. 2004. How market smarts can protect property rights. *Harvard Bus. Rev.*, **82**(12) 72.
- Arora, A., A. Fosfuri, A. Gambardella. 2001. *Markets for Technology: The Economics of Innovation and Corporate Strategy*. The MIT Press, Cambridge.
- Audretsch, D.B., M.B. Feldman. 1996. R&D spillovers and the geography of innovation and production. *Amer. Econom. Rev.* **86** 630–640.
- Bartlett, C.A. S. Ghoshal. 1990. The multinational corporation as an inter-organizational network. *Acad. of Management Rev.* **15** 603–625.
- Baum, J.A.C., H.A. Haveman. 1997. Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898–1990. *Admin. Sci. Quart.* **42**(2) 304–338.
- Baum, J.A.C., O. Sorenson, eds. 2003. Geography and strategy. *Advances in Strategic Management, Vol. 20*. JAI/Elsevier, Oxford.
- Buckley, P.J., C.M. Casson. 1976. *The Future of the Multinational Enterprises*. Macmillan: London.
- Buenstorf, G., S. Klepper. 2004. The origin and location of entrants in the evolution of the U.S. tire industry. *Papers in Economics and Evolution 2004-07*. Max Planck Institute of Economics.
- Cassiman, B., R. Veugelers. 2002. R&D cooperation and spillovers: Some empirical evidence from Belgium. *Amer. Econom. Rev.* **92**(4) 1169–1184.
- Chung, W., J. Alcácer. 2002. Knowledge seeking and location choice of foreign direct investment in the United States. *Management Sci.* **48** 1534–1554.
- Cohen, W.M., D. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Quart.* **24** 128–152.

- Cohen, W.M., R.R. Nelson, J.P. Walsh. 2000. Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). *NBER Working Paper* 7552.
- Cohen, W.M. 1995. Empirical studies of innovative activity. In *Handbook of the Economics of Innovations and Technological Change*, Paul Stoneman, ed. Wiley-Blackwell, Oxford.
- Cummings, J.N. 2004. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Sci.* **50**(3) 352–364.
- Dierickx, I., K. Cool. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Sci.* **35** (12) 1504–1511.
- Edstrom, A., J.R. Galbraith. 1977. Transfer of managers as a coordination and control strategy in multinational organizations. *Admin. Sci. Quart.* **22**(2) 248-263.
- Feinberg, S., A. Gupta. 2009. MNC subsidiaries and country risk: Internalization as a safeguard against weak external institutions. *Acad. of Management J.* **52**(2), in press.
- Fleming, L., C. King III, A. Juda. 2006. Small worlds and regional innovation. *Organization Sci.*, forthcoming.
- Furman, J.L., M.K. Kyle, I. Cockburn, R. Henderson. 2006. Public and private spillovers, location, and the productivity of pharmaceutical research. *Annales d'Economie et Statistique* **79/80**.
- Gilson, R.J. 1999. The legal infrastructure of high technology districts: Silicon Valley, Route 128, and covenants not to compete. *New York University Law Rev.* **74** 575–629.
- Gupta, A.K, V. Govindarajan. 2000. Knowledge flows within multinational corporations. *Strategic Management J.* **21** 473-496.
- Hall, B.H., A.B. Jaffe, M. Trajtenberg. 2005. Market value and patent citations. *Rand J. of Econom.* **36**(1) 16–38.
- Helfat, C.E., 1994. Firm-specificity in corporate applied R&D. *Organization Sci.* **5** 173–184.
- Henisz, W., J. Macher. 2004. Firm- and country-level tradeoffs and contingencies in the evaluation of foreign investment: The semiconductor industry, 1994–2002. *Organization Sci.* **15**(6) 537–554.
- Jacobs, J. 1969. *The Economy of Cities*. Random House, New York.

- Jaffe, A.B. 1986. Technological opportunity and spillovers of R and D: evidence from firms' patents, profits, and market value. *Amer. Econom. Rev.* **76**(5) 984–1001.
- Jaffe, A.B., M. Trajtenberg, R. Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quart. J. of Econom.* **434** 578–598.
- Jaffe, A.B., M. Trajtenberg, M.S. Fogarty. 2000. The meaning of patent citations: report on the NBER/Case-Western Reserve Survey of patentees. *NBER Working Paper* 7631.
- Kogut, B., U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Sci.* **3**(3) 383–397.
- Kogut, B., U. Zander. 1993. Knowledge of the firm and the evolutionary theory of the multinational corporation. *J. Intern. Bus. Stud.* **24** 625–645.
- Gittelman, M., B. Kogut. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citations patterns. *Management Sci.* **49**(4) 366–382.
- Lahiri, N. 2003. Knowledge spillovers: geography, technology and intra-firm linkages. Doctoral dissertation: Ross School of Business, University of Michigan.
- Liebeskind, J.P. 1996. Knowledge, strategy, and the theory of the firm. *Strategic Management J.* **17**(Winter Special Issue) 93–109.
- Levin, R.C., A.K. Klevorick, R.R. Nelson, S.G. Winter. 1987. Appropriating the returns from industrial R&D. *Brookings Papers on Economic Activity* **3** 783–820.
- Mansfield, E., M. Schwartz, S. Wagner. 1981. Imitation costs and patents: An empirical study. *Economic J.* **91** 907–918.
- Marshall, A. 1922. *Principles of Economics*, 8th edition. MacMillan, London.
- Nobel, R., J. Birkinshaw. 1998. Innovation in multinational corporations: control and communication patterns in international R&D operations. *Strategic Management J.* **19**(5) 479–496.
- Porter, M.E. 1998. *On Competition*. Harvard Business School Press, Cambridge, MA.
- Rajan, R., L. Zingales. 2001. The firm as a dedicated hierarchy: A theory of the origins and growth of firms. *Quart. J. of Econom.* **116** (3) 805–851.

- Sampat, B. 2009. When Do Applicants Search for Prior Art.? *J. of Law and Econom.* Forthcoming.
- Sanna-Randaccio, F., R. Veugelers. 2002. Multinational knowledge spillovers with centralized versus decentralized R&D: A game theoretic approach. *CEPR Discussion Paper* 3151.
- Saxenian, A. 1996. *Regional Advantage*. Harvard University Press, Cambridge, MA.
- Shaver, J.M., F. Flyer. 2000. Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management J.* **21** 1175–1193.
- Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Sci.* **51**(5) 756–770.
- Swaminathan, A., J. Delacroix. 1991. Differentiation within an organizational population: Additional evidence from the wine industry. *Acad. of Management J.* **34**(3) 679–692.
- Stuart, T.E. 2000. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic Management J.* **21**(8) 791–811.
- Stuart, T.E., J.Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management J.* **17** 21–38.
- Stuart, T.E., O. Sorenson. 2003. Liquidity events and the geographic distribution of entrepreneurial activity. *Admin. Sci. Quart.* **48** 175–201.
- Szulanski, G. 1996. Exploring internal stickiness: impediments to the transfer of best practices within the firm. *Strategic Management J.* **17** 27–43.
- Teece, D. 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy* **15**(6) 285–305.
- Teece, D.J., G. Pisano, A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management J.* **18**(7) 507-533.
- Trajtenberg, M., R. Henderson, A. Jaffe. 1997. University versus corporate patents: a window on the basicness of invention. *Economics of Innovation and New Technology* **5**(1) 19–50. Also in *Patents, citations, and innovations*, 2002, ed. A. Jaffe, M. Trajtenberg, p. 51-88. MIT Press, Cambridge, MA.

- Vickers, J. 1986. The evolution of market structure when there is a sequence of innovations. *J. of Industrial Econom.* **35**(1) 1–12.
- Yoffie, D.B. 1993. Foreign direct investment in semiconductors. *Foreign Direct Investment*, ed. K.A. Froot, p. 197–230. University of Chicago Press, Chicago.
- Zhao, M. 2006. Conducting R&D in countries with weak intellectual property rights protection. *Management Sci.* **52**(8) 1185–1199.
- Ziedonis, R. H. 2004. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management Sci.* **50**(6) 804–820.

Table 1. Descriptive Statistics

	Obs.	Mean	St Dev	Min.	Max.
Dependent variables					
self_citations (inventor)	5,266	1.09	3.20	0	57
self_citations (inventor+examiner)	10,204	1.05	2.79	0	63
cross_cluster	1,089	4.42	9.23	0	90
to_primary	1,089	1.65	4.10	0	57
Independent variables					
Competition based on technology*					
innovators	304	5.42	13.58	1	130
innovators_profit	304	4.92	12.40	1	124
small_innovators	304	0.72	2.98	0	33
large_innovators	304	4.20	9.95	1	101
innovators_nonprofit	304	0.50	1.44	0	12
universities	304	0.29	0.94	0	8
govt_innovators	304	0.15	0.59	0	6
other_nonprofit	304	0.06	0.24	0	2
Competition based on product market*					
in_industry	304	1.19	3.75	0	45
Not_in_industry	304	4.14	10.46	1	92
in_segment	304	1.28	3.48	0	38
Not_in_segment	304	4.23	10.14	1	85
competitors	304	0.81	1.93	0	14
Not_competitors	304	4.61	12.13	1	117
Cluster variables*					
Plants_in_cluster	304	2.54	7.92	0	75
fabless_in_cluster	304	1.67	13.00	0	211
publications_in_cluster	304	21.04	46.31	0	515
Firm-cluster variables*					
with_plant	304	0.05	0.21	0	1
with_fabless	304	0.00	0.00	0	0
Firm variables*					
Plants	16	18.03	8.11	7	36
fabless	16	0.19	0.40	0	1
patents_total	16	1,295.43	589.72	516	2,702
patents_semi	16	199.97	102.96	120	530

* Statistics are based on averages across the years 1998-2000

Table 2. Negative Binomial estimates on self-citations

Dependent variable: Self-citations; Exposure variable Total Citations						
	(1)	(2)	(3)	(4)	(5)	(6)
plants_in_cluster	-0.005 (1.71)†	-0.006 (1.74)†	-0.014 (3.73)**	-0.010 (2.75)**	-0.006 (1.89)†	-0.007 (2.12)*
fabless_in_cluster	-0.001 (0.77)	-0.001 (0.85)	-0.002 (1.90)†	-0.003 (2.07)*	-0.001 (0.94)	-0.001 (0.85)
publications_in_cluster	0.000 (0.58)	0.000 (0.68)	-0.001 (1.13)	-0.000 (0.18)	0.000 (0.77)	0.000 (0.27)
with_plant	0.132 (2.99)**	0.130 (2.92)**	0.185 (3.97)**	0.169 (3.58)**	0.121 (2.70)**	0.134 (2.99)**
with_fabless	-0.053 (0.19)	-0.051 (0.18)	-0.159 (0.58)	-0.111 (0.40)	-0.159 (0.56)	-0.077 (0.28)
patents_total	0.000 (0.18)	0.000 (0.17)	0.000 (0.68)	0.000 (0.19)	0.000 (0.25)	0.000 (0.15)
patents_semi	0.002 (7.71)**	0.002 (7.70)**	0.002 (8.13)**	0.002 (7.73)**	0.002 (7.82)**	0.002 (7.81)**
Innovators	0.001 (1.05)					
innovators_profit		0.002 (1.01)				
small_innovators			-0.012 (3.53)**			
large_innovators			0.015 (4.47)**			
in_industry				0.028 (2.78)**		
not_in_industry				-0.001 (0.25)		
in_segment					0.006 (1.87)†	
not_in_segment					0.0007 (0.40)	
competitors						0.015 (2.06)*
not_competitors						0.0008 (0.47)
innovators_nonprofit		-0.003 (0.25)		-0.003 (0.26)	-0.001 (0.10)	0.002 (0.16)
Universities			-0.009 (0.76)			
govt_innovators			0.006 (0.18)			
other_nonprofit			0.103 (1.40)			
Constant	-17.434 (0.01)	-16.192 (0.03)	-16.903 (0.02)	-17.420 (0.01)	-17.404 (0.01)	-16.167 (0.03)
Firm fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Country fixed effects	Y	Y	Y	Y	Y	Y
Observations	5,117	5,117	5,117	5,117	5,117	5,117
Log Likelihood	-5776.87	-5776.8	-5765.12	-5773.3	-5775.57	-5775.05

Absolute value of z statistics in parentheses

† significant at 10%; * significant at 5%; ** significant at 1%

Table 3. Mean tests for citations and self-citations in different competitive environments

Panel A: Comparing average citations across quartiles of competitive environments

		Quartiles for competitive environment ^		
		1	2	3
Ave. citations		3.57	3.25	3.40
Quartiles for competitive environment ^	2	3.25	-0.320 0.367	
	3	3.40	-0.168 0.378	0.152 0.376
	4	3.20	-0.371 0.388	-0.051 0.386

Panel B: Comparing average self-citations across quartiles of competitive environments

		Quartiles for competitive environment ^		
		1	2	3
Ave. citations		1.72	1.09	0.73
Quartiles for competitive environment ^	2	1.09	-0.625 * 0.234	
	3	0.73	-0.988 * 0.229	-0.363 * 0.250
	4	0.52	-1.200 * 0.241	-0.575 * 0.261

First value corresponds to (row mean-column mean), second value is t-statistic

* 5% significance level

^ Competitive environment measured based on direct competitor data from Hoovers. Quantile 1 corresponds to clusters with the largest number of competitors, quantile 4 to clusters with the smallest number of competitors.

Table 4. Citations from competitors: Within clusters vs. outside of clusters

	Number of citing patents from various entities			Ratio of competitor citations over citations of all firms		
	Outside cluster	Within cluster	Total	Outside cluster	Within cluster	Total
Universities	222	3	225			
Other non-profit organizations	14	3	17			
Government	19	1	20			
Competitors in industry	10,789	830	11,619	1.47	0.87	0.58
Other firms	7,357	958	8,315			
Total firms	18,401	1,795	20,196			
Competitors in segment	12,196	1,105	13,301	2.05	1.62	0.67
Other firms	5,950	683	6,633			
Total firms	18,146	1,788	19,934			
Competitors	12,275	1,107	13,382	2.09	1.63	0.67
Other firms	5,871	681	6,552			
Total firms	18,146	1,788	19,934			

Table 5. Negative Binomial estimates on cross-cluster links

Dependent variable Exposure variable	Patents with cross-cluster links Total number of patents					Cross-cluster links with primary clusters All cross-cluster links				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
plants_in_cluster	-0.019 (4.48)**	-0.019 (4.30)**	-0.023 (5.23)**	-0.020 (4.67)**	-0.023 (5.36)**	-0.027 (4.11)**	-0.034 (4.44)**	-0.033 (4.77)**	-0.028 (4.26)**	-0.030 (4.50)**
fabless_in_cluster	-0.001 (0.74)	-0.001 (0.55)	-0.002 (1.34)	0.000 (0.06)	-0.001 (0.48)	0.0017 (0.66)	0.0011 (0.43)	-0.0003 (0.11)	0.0046 (1.82)†	0.0011 (0.43)
publications_in_cluster	0.002 (2.93)**	0.002 (2.34)*	0.001 (2.06)*	0.002 (3.06)**	0.001 (2.14)*	0.003 (3.15)**	0.002 (1.65)†	0.002 (2.38)*	0.003 (3.46)**	0.003 (2.83)**
with_plant	0.783 (9.94)**	0.782 (9.94)**	0.789 (10.12)**	0.797 (10.07)**	0.773 (9.96)**	0.783 (7.49)**	0.792 (7.59)**	0.780 (7.47)**	0.808 (7.77)**	0.769 (7.38)**
with_fabless	-0.064 (0.28)	-0.053 (0.23)	-0.045 (0.20)	-0.056 (0.24)	-0.044 (0.19)	0.580 (1.50)	0.622 (1.61)	0.611 (1.57)	0.819 (2.10)*	0.658 (1.69)†
patents_total	-0.001 (2.90)**	-0.001 (3.04)**	-0.001 (3.01)**	-0.001 (2.94)**	-0.001 (2.80)**	-0.001 (2.40)*	-0.002 (2.58)**	-0.002 (2.47)*	-0.001 (1.91)†	-0.001 (2.41)*
patents_semi	0.000 (1.34)	0.000 (1.36)	0.000 (1.15)	0.000 (1.38)	0.000 (1.50)	0.000 (2.04)*	0.000 (1.95)†	0.000 (1.88)†	0.000 (2.23)*	0.000 (2.18)*
innovators_profit	0.006 (2.18)*					0.005 (1.37)				
small_innovators		0.002 (0.26)					-0.010 (1.23)			
large_innovators		0.007 (1.89)†					0.018 (2.59)**			
in_industry			0.038 (4.06)**					0.044 (3.30)**		
not_in_industry			0.000 (0.05)					0.000 (0.03)		
in_segment				0.009 (2.69)**					0.017 (3.72)**	
not_in_segment				-0.005 (0.67)					-0.029 (3.10)**	
Competitors					0.046 (5.00)**					0.034 (2.58)**
not_competitors					0.002 (0.91)					0.004 (1.02)
innovators_nonprofit	0.000 (0.03)		0.011 (0.70)	-0.006 (0.37)	0.007 (0.48)	-0.011 (0.43)		-0.001 (0.03)	-0.040 (1.55)	-0.010 (0.39)
Universities		-0.003 (0.18)					-0.031 (1.06)			
govt_innovators		0.044 (1.34)					0.080 (1.68)†			
other_nonprofit		-0.038 (0.51)					-0.0001 0.00			
primary_cluster	1.44 (16.13)**	1.45 (16.15)**	1.44 (16.29)**	1.45 (16.19)**	1.46 (16.50)**					
Constant	-0.333 (1.09)	-0.309 (1.01)	-0.319 (1.05)	-0.335 (1.10)	-0.258 (0.85)	-0.197 (0.50)	-0.164 (0.42)	-0.175 (0.45)	-0.215 (0.55)	-0.127 (0.32)
Firm fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Country fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,030	1,030	1,030	1,030	1,030	966	966	966	966	966
Log Likelihood	-2256.02	-2254.66	-2249.54	-2254.95	-2245.82	-1529.49	-1525.79	-1524.96	-1523.95	-1526.85

Absolute value of z statistics in

† significant at 10%; * significant at 5%; ** significant at 1%