Global Competitors as Next-Door Neighbors:

Competition and Geographic Co-location in the Semiconductor Industry

Juan Alcácer
Harvard Business School
Soldiers Field, Morgan Hall 227
(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 2008

Global Competitors as Next-Door Neighbors:

Competition and Geographic Co-location in the Semiconductor Industry

ABSTRACT

Despite the many advantages offered by technology clusters, firms located in them face the risk of losing valuable knowledge to nearby competitors. In this study, we argue that multi-location firms strategically organize their R&D activities to appropriate the value of innovations generated in clusters, mainly through three mechanisms: technological distance, value internalization, and control. Empirical analysis of the global semiconductor industry provides supportive evidence of such mechanisms. In clusters where direct competitors are right next door, leading firms generate innovations that are technologically distant from their neighbors, have more internalized value, and involve inventors from other geographic locations, particularly from headquarters. Interestingly, the strategies seem to be much more sensitive to neighboring firms competing in the same marketplace than those sharing the same technological space. The findings offer important insights into the interaction between firms' internal organization and their external environment.

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

We acknowledge the helpful comments of Wilbur Chung, Heather Berry, Lori Rosenkopf, Myles Shaver, Nicolaj Siggelkow, Sid Winter, and Rosemarie Ziedonis as well as seminar participants at Carnegie Mellon, Duke, Harvard, Michigan, NYU and Wharton. Errors remain our own.

1. Introduction

Technology clusters – geographic concentrations of firms and other institutions engaging in closely related R&D activities – are characterized by intensive local knowledge flows across organizational boundaries (Porter 2000). Due to the tacit nature of knowledge, 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, an additional mechanism of knowledge flow across organizations.

Knowledge, however, flows in both directions. Unintended knowledge outflows to competitors can erode the competitive advantage held by industry leaders, a potential cost that may overweigh the many benefits of clustering. Consequently, previous research suggests that leading firms may choose to locate apart from technology clusters, using geographic distance as a strategy to preserve and appropriate value from innovation (e.g., Shaver and Flyer 2000). Nevertheless, this approach may be neither desirable nor sustainable. Leading firms are often attracted to a technology cluster by some unique advantages the location has to offer: proximity to research universities, favorable government policies, abundance of human capital, etc. Even when a leading firm decides to locate apart, it has little control over subsequent location decisions by competitors. Hence, voluntary or not, co-location with competitors happens, and most innovation in technology-driven industries – including that of industry leaders – still occurs overwhelmingly in technology clusters.

What enables the leading firms to benefit from the location-specific advantages without compromising their ability to profit from innovation? In this paper, we address this question by integrating two important factors in cluster dynamics: geographically dispersed R&D organization at the firm level, and the composition of various organizations – competitors or not – at the cluster level.

At the firm level, large multi-location firms are dynamically mobilizing and integrating knowledge on a global basis (Bartlett and Ghoshal 1990). To understand R&D dynamics in a cluster, we recognize that an entity in the cluster may well be part of an extended organization, with its multiple locations and multiple business lines strategically integrated. 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.

At the cluster level, previous studies often overlooked the multi-dimensional relationships among the local entities (Cohen 1995: p.230). Firms in a technological cluster may share similar technological backgrounds or even engage in patent races against each other, but they do not necessarily compete in the same product market. Industry-specific market information and other complementary resources can significantly reduce the concern over knowledge exchanges, thus allowing symbiotic relationships to develop in technology clusters.

We argue that multi-location firms that face heterogeneous local competitive environments can benefit from technology clusters while minimizing potential information leakages through three mechanisms. First, a firm may allocate technologies that do not overlap much with those of nearby competitors to technology clusters, using technological distance to minimize information leakage. Second, a firm may choose to develop technologies in clusters that can be quickly built on by other units of the firm, thus reducing imitation incentives of nearby firms. Finally, a firm may simply exert tighter control over local R&D by increasing the involvement of researchers from other locations, particularly from headquarters. These are all strategies available to geographically dispersed organizations.

Examining the semiconductor industry from 1998 to 2001, we find supportive evidence at both the firm and the location level. When surrounded by direct competitors, the technology leaders in the industry are likely to keep a larger technological distance from local entities, develop technologies that are cited more internally, and organize more cross-regional collaborations. Moreover, the three mechanisms are highly responsive to competition when competition is defined as firms operating in the same product market, but are much less visible when competition is defined as firms sharing the same knowledge base.

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. Firm Strategy and Technological Clusters

2.1 Appropriability through Geographic Distance

There has been a revival of interest in the geographic concentration of economic activities (Porter 2000) and its implications for firms' strategic location choices (Chung and Alcácer 2002, Baum and Sorenson 2003). Starting with the seminal work by Marshall (1920), researches have suggested that firms in an industry cluster benefit from inter-firm knowledge spillover, access to specialized labor, and access to specialized intermediate inputs. Among the various activities along the value chain, R&D activities depend on knowledge spillover the most and thus show the highest level of concentration (Audretsch and Feldman 1996, Alcácer 2006). In particular, geographic proximity in technology clusters 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.

However, inter-firm knowledge spillovers may also hinder a firm's ability to appropriate value from its own innovations. For example, in the context of collaborative R&D, Cassiman and Veugelers (2002) distinguish incoming spillover, which enhances innovation, from appropriability, the ability to generate economic rents from intellectual capital. In the context of strategic alliances, appropriability issues are often the cause of concerns for alliance partners (Oxley and Sampson 2004).

Similar arguments have been raised for firms' location decisions. Shaver and Flyer (2000) suggest that firms not only benefit from the knowledge flows in clusters, but also contribute to them. Thus, firms locate apart when the costs of clustering exceed the benefits. Linking knowledge spillover to strategic location choices, Alcácer and Chung (2006) suggest that, because knowledge flows in both directions, industry leaders may shy away from clusters. Earlier, Yoffie (1993) also notes that semiconductor managers avoid locating near competitors for fear of technology spillover to other firms.

Minimizing knowledge outflow through geographic distance is the common thread in these papers. However, some caveats remain, in both theory and empirics. Theoretically, geographic distance may not be a sustainable strategy over time. Even if a leading firm decides to locate apart, it has little control over the subsequent location decisions of competitors, who may relocate or open new facilities nearby to take advantage of positive externalities. In addition, new firms may emerge as previous employees start up new businesses locally (Buenstorf and Klepper 2004). To the extent that industry leaders cannot stop other firms from collocating, geographic distance offers only a temporary measure to prevent unintended knowledge outflows.

Empirically, most studies examine individual location choices (i.e., marginal entries), ignoring the multiple locations that a firm has already had presence in. Micron Technology, Inc., for example, has R&D presence in Silicon Valley, but it also maintains a strong R&D and manufacturing base in Boise, Idaho, where it owns over 98% of the locally developed semiconductor patents. That is, firms' choices at one location should be analyzed in combination with the overall organizational structure. Finally, the evidence of leading firms locating apart is often industry-dependent. When key inputs, including specialized knowledge, are only available in a specific location, even industry leaders flock there. For example, Chung and Alcácer (2002) find that even the most technologically advanced foreign firms locate next to existing firms in some technology-intensive industries, such as semiconductors and pharmaceuticals.

If staying away is not an option, how do industry leaders manage R&D in technology clusters so that they can benefit from the local resources (e.g., human capital, tax incentives and university facilities) while still appropriating value from innovation? In the rest of the section, we propose three mechanisms that a multi-unit multi-location firm may utilize to maintain its competitive edge: keeping certain technological distance from local competitors, developing innovations with high internalized value, and maintaining tight control over R&D activities in technology clusters.

2.2 Appropriability through Technological Distance

In essence, geographic distance is one of the many modes of strategic differentiation, by which firms set themselves apart from rivals to soften competition and obtain higher returns. Other strategies of differentiation have been widely discussed in strategy (Porter 1985) as well as organizational theory (Baum and Haveman 1997). For example, in a study of Californian wineries between 1940 and 1984, Swaminathan and Delacroix (1991) find that wineries managed to escape competitive pressure through organizational differentiation; those that entered the table wine niche were subject to lower failure rates.

In clustered areas where geographic distance is small, technological distance – engaging in R&D that is technologically differentiated from that conducted by nearby rivals – may serve as an alternative mechanism for knowledge protection. For any knowledge spillover to occur, the receiving firm must be able to recognize and absorb the specific knowledge. Because organizations tend to search for knowledge around their own technological positions (Stuart and Podolny 1996), certain technological distance may help the innovating firms to stay off rivals' radar screen. Moreover, even if the technologies are recognized by a rival, the ability to identify, assimilate, and apply new technologies is highly dependent on the rival's prior knowledge (Cohen and Levinthal 1990). Therefore, leading firms may reduce the imitation concern by developing technologies that deviate from key competitors' local R&D portfolios. *We hypothesize that leading firms are more likely to distance themselves technologically from the local competitors as the number of competitors in the cluster increases*.

Furthermore, the level of R&D activity performed by competitors in the cluster may mediate the above relationship. For example, a firm may perceive a lower risk of knowledge outflow if competitors only conduct limited amount R&D in the cluster. Because a firm's absorptive capacity depends on its previous stock of knowledge and knowledge is locally bounded, a competitor's ability to receive knowledge should be the highest in its core cluster – the cluster where it performs most of its R&D. Therefore, *we expect to observe larger technological distance in clusters that are core clusters for a competitor*.

Admittedly, managing technological distance comes with costs. First, technological distance may also affect a firm's own absorptive capacities and prevent it from taking full advantage of the local knowledge spillover. However, to the extent that the cluster is populated by organizations playing different roles along the value chain or belonging to different industries, a firm can still benefit from knowledge developed in the cluster while keeping appropriability of its own technologies (Cohen 1995). Such clusters, defined by Jacobs (1969) as locations where firms share a broad technological domain but do not compete on the same product markets, extend Marshall's concept of a cluster to reflect the emergence of cross-industry agglomerations. Firms in Jacobian clusters can still benefit from local talent pool or specialized suppliers, emerging from a common technology, while minimizing knowledge outflows to direct competitors. Second, the purpose of carrying out R&D is to serve the firm's strategic goals, which are established for the long run and are not easily altered. In other words, the type of R&D needed for the firm's long-term development is, to a large extent, predetermined. Fortunately, this is less of a problem for multi-location firms, who usually manage large portfolios of R&D projects. They are able to allocate projects strategically to specific locations without altering the overall innovation objectives.

2.3 Appropriability through Internalization

Knowledge assets are not only interconnected, but also cumulative in nature (Dierickx and Cool 1989). Most commercial offerings are the result of a long sequence of technological improvements (Vickers 1986). If a firm can build on its new technologies more efficiently than potential imitators, then it may gain the crucial lead time in the marketplace and still appropriate value from innovation.

In fact, the innovating firms are often in an advantageous position to identify and further develop the locally produced intellectual properties, despite the intensive information flow in technology clusters. First, due to the high uncertainty and tacitness associated with technologies, 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 move on to the next stage without the costly learning process (Mansfield *et.al*. 1981). Second, for a geographically dispersed organization, knowledge that is difficult to codify or teach can be more efficiently transferred within the firm. Technologies developed in one technology cluster can be transferred internally, facilitating the accumulation and integration of knowledge throughout the organization (Bartlett and Ghoshal 1990, Kogut and Zander 1993).

Internalization as a mechanism of value appropriation has also been tested in various empirical settings. For example, Zhao (2006) suggests that the ability to integrate and build on internal technologies enables multinational firms to conduct R&D in countries with weak intellectual property rights protection. There, firms develop technologies that are intricately related to the firms' expertise residing elsewhere in the world, so local imitation of these specific technologies does not affect the firms' value appropriation on the global market. Similarly, based on a large sample of U.S. headquartered multinational firms, Feinberg and Gupta (2006) find strong evidence that firms respond to high risks in the host countries by increasing the extent of internal transactions. Thus, we argue that firms can appropriate value from their R&D in technology clusters if they can internalize the resultant technologies better and faster than nearby competitors, and the importance of such strategies increases with the intensity of local competition. Accordingly, *we expect higher levels of internalization for technologies developed in clusters with larger numbers of competitors*.

2.4 Appropriability through Control

When facing appropriability risks, firms also strategically adjust their organization and governance structures for the protection of valuable intellectual properties. In the strategic alliance literature, firms are found to adopt more hierarchical, instead of contractual, governance modes when knowledge protection is weak (Oxley 1999), and that they carefully limit the scope of knowledge sharing when partnering firms are direct competitors (Oxley and Sampson 2004).

From a theoretical perspective, Rajan and Zingales (2001) 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 critical resources, enforcing property rights is difficult, and competition is fierce. In such circumstances, controlling the access to certain key resources renders each division alone less efficient than the whole, thus preventing the risk of expropriation. Similarly, Liebeskind (1996) argues that disaggregating tasks gives firms an advantage in knowledge protection, especially when reinforced by spatial isolation. In other words, appropriability risk is reduced if each division is highly dependent on its linkage to other parts of the organization.

One indication of such linkages is the participation of researchers from other locations in local R&D projects (Nobel and Birkinshaw 1998). Edstrom and Galbraith (1977) propose that transferring managers across units can be considered a coordination and control mechanism in multinational organizations. Based on knowledge specialization and interpersonal networks, this kind of rotation facilitates control at the corporate level while still allowing autonomy and flexibility at the local level – features that are important for innovative activities. Given the nontrivial costs of organizing such cross-regional linkages (Cummings 2004), *we would expect more internal linkages in regions with higher appropriability risks, e.g. in clusters with a large number of direct competitors*.

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 of success in semiconductors. Firms invest relentlessly in R&D to introduce new products and improve production processes (Stuart 2000). Moreover, firms in the industry routinely patent their innovations, and patent data have been used to trace the traits and geographic distribution of innovation. Second, the benefit of cross-firm knowledge spillover has been documented as the driver of agglomeration in the industry (Saxenian 1994, Fleming, *et al.* 2006). The high levels of geographic concentration in this industry also suggest that firms may have already developed appropriability strategies to manage outward spillover. Third, this is a truly global industry with leading firms operating at multiple locations around the world, and there exists significant heterogeneity among the semiconductor firms in terms of their product markets, R&D portfolios, position in the value chain and geographic locations. Firms range from industry giants that participate in activities throughout the value chain to enterprises that specialize in design (known as fabless) or testing, 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 different sources. First, we identify innovating firms using patent data from the *Derwent World Patent Index* (DWPI), a well-recognized dataset that encompasses over 30 million patent documents from 41 patent-issuing authorities worldwide, and we rely on Derwent's technological classification[1] 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 duplications. Thus, we 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, locations, and differ only in the scope of their claims. The final sample consists of 23,675 patent families[2] whose assignees are American and foreign firms, universities, as well as government- and industry-sponsored research labs. For the 300 patent families that have more than one assignee, all assignees, 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%).

---

[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/dwpiref/reftools/classification.

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

This initial sample is supplemented with directories of semiconductor plants, fabless companies, and institutions behind scientific publications. Information on 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 over 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 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 the local environments, our analysis of R&D strategies is focused on the top 16 innovating firms in the industry[3]. These large multinational firms correspond to the top 1% of all firms in terms of patent families, representing 50% of the patent output

[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.

and 40% of the plants operating in this period. 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% firms in terms of patent families – and obtain similar results.

3.2 Cluster Definition

Defining technology clusters is a crucial element in our empirical setup. Instead of relying on predetermined administrative boundaries, such as states or metropolitan areas, we apply a mathematical algorithm that uses latitude and longitude data to identify technological clusters. Two main reasons justify this decision. First, there is not an administrative unit that is universally 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 predetermined administrative boundaries, which is clear after a quick inspection of inventor locations in, e.g., northeast U.S. and central Japan. One administrative unit may encompass multiple clusters, while one technological cluster may expand across administrative lines.

In this study, we define clusters based on the actual distribution of inventor locations, following a three-step approach. First, we identify the location of each element (i.e., a patent inventor, plant, fabless company or scientific publication) in the sample, and match the locations to two comprehensive sources of geographic names. For U.S. locations, we use the *Geographic Names Information System* (GNIS) of the U.S. Geological Survey and obtain latitude and longitude information for all 38,261 locations in the country. 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 also includes phonetic variations for the spellings from a different alphabet (e.g., Asian countries) or an alphabet with extra characters (e.g., Scandinavian and Slavic countries). Ambiguous matches are checked manually by native speakers 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 many other traditional clustering methods do – 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 are not far away from each other. 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, fabless 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 miles[4]. 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 fabless and plant data, respectively[5].

3.3 Dependent Variables

3.3.1 Technological Distance

Following Jaffe (1986), we construct technological vectors at the firm-location-year level to measure the technological distance among firms. The vectors are built based on Derwent's technological classification system. The technological distance between two vectors i and j is defined as:

---

[4] We also tried other minimum distances, e.g., 20, 25 and 30 miles, with very similar outcomes.

[5] Note that 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.

$$D_{ij} = 1 - \frac{v_i v_j'}{\sqrt{(v_i v_i')(v_j v_j')}} \in (0,1) \qquad (1)$$

A feature of this measure, as Jaffe (1986) points out, is that the absolute number of patents does not affect the result; what matters is the structural distribution across technology classes. $D_{ij} = 1$ when the two vectors are orthogonal (i.e. firms' innovations have no overlap at all) and $D_{ij} = 0$ when they are parallel (i.e. firms' innovations fully overlap in the technological space). To determine whether firm *i* chooses to keep certain technological distance from other players, we calculate pair-wise technological distance D_{ij} between the local patents developed by the focal firm *i* and those developed by every other firm *j* that has R&D in the same cluster-year. The average of all these pair-wise measures – $technological\_distance_{ict}$ – is used as firm *i*'s technological distance from local innovators of cluster *c* in year *t*.

3.3.2 Internalized Value

A key concept in this study is the extent to which the value of an innovation is appropriated by 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 that belong to the same firm represent internalized knowledge transfers leading to the firm's competitive advantage. Hence, we use forward self-citations as a proxy for the internalized value of technologies. 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 is the sum of citations to 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 2006). To avoid this problem, our main models are estimated using citations listed by inventors only. In our sample, about 38% of the patent families that receive at least one inventor citation also have at least one self-citation. The number is 30% when both inventor and examiner citations are considered.

3.3.3 Control

Geographically decentralized R&D in a multi-location firm increases the challenges of effectively appropriating returns from innovation and preventing outward spillover to competitors (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' home bases – locations where firms conduct most of their R&D – from their peripheral R&D facilities. We then identify the cross-cluster links that involve the home bases and those connecting peripheral locations only. As shown in Table 1, the average number of cross-cluster links per firm-cluster-year is 4.4. Cross-cluster teams involving the home bases are more common than teams that link 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* capture the number of for-profit and nonprofit assignees, respectively. In addition, we use the status information on patent applications[6] 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:

[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.

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 semiconductors[7], as well as the names of its direct competitors. Then we count the number of for-profit assignees that are 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*).

We complete the characterization of local innovation environments with three more variables: *plants_in_cluster*, *fabless_in_cluster* and *publications_in_cluster*, which represent the numbers of plants, fabless companies and publications per cluster-year. In addition, we use two dummy variables, *with_plant* and *with_fabless*, to indicate whether a particular firm has plants or fabless 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 up to 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.

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 three dependent variables – *technological_distance*, *self_citation*, and *cross_cluster* – are the three dimensions that characterize local innovations and are correspondent to the three appropriability strategies described in the previous section.

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

$$DV_{ict} = C_{ict} + X_{ict} + Y_{ct} + \zeta_t + \upsilon_i + \tau_{ctry} + \varepsilon_{\mathrm{ict}} \qquad (2)$$

[7] Hoover's reported 13 segments under semiconductors, including memory chips & modules, microprocessors, etc.

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 υ_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 analyses for *technological_distance* and *cross_cluster* are conducted at the firm-cluster level. As *self_citation* and *cross_cluster* are both count variables, negative binomial models are used for the estimations[8]. For *technological_distance*, a continuous variable between 0 and 1, we use both OLS and Tobit model for the estimation and obtain consistent results. Only the OLS results are presented due to space constraints.

4. Empirical Results

4.1 Analysis of Technological Distance

Table 2 presents the estimates for *technological_distance* using OLS and the DWPI classification system[9]. Recall that the distance measure is equal to 0 when two technology vectors overlap and 1 when they are orthogonal. Therefore, positive coefficients indicate a divergence in technology (a larger technological distance) and negative coefficients indicate a convergence (a smaller distance).

Our prior is that firms diverge from technologies developed in the local community when the imitation threat from nearby competitors is high. As shown in models (1) and (2), the focal firm develops innovations that are more technologically distant as the number of for-profit innovators, particularly large for-profit innovators, increases. In models (3) to (5), we measure the competitive environments following the product-market definition. Across specifications, positive and highly significant coefficients of *in_industry, in_segment* and *competitors* indicate that the presence of competing firms induces larger

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

9 We also construct the technology vectors based on IPC or USC, and obtain similar, but statistically weaker results.

technological distance. Interestingly, a larger presence of nonprofit innovators is associated with technological proximity. A potential explanation for this finding is that firms feel less threatened by nonprofit organizations and are more attracted by the learning opportunities that they offer.

Regarding our control variables, the coefficient of *with_plant* is positive and significant. This may be caused by innovations that are closely linked to the manufacturing processes, which tend to be more firm-specific. Overall, the findings are consistent with the argument that firms keep a larger technological distance from local innovations when there is more direct competition in the neighborhood.

Some caveats may arise from using the average pair-wise distance as dependent variable. For example, a firm that keeps a large technological distance from direct competitors but is close to any other organizations in the cluster may end up with a small average distance. Moreover, a new entrant with its technological niche may inadvertently increase the average distance for every firm in the cluster. To overcome this issue, we conduct further analysis at the dyad level, using the dyadic technological distance between two organizations in a cluster. Specifically, we estimate the equation:

$$technological\_distance_{ijct} = comp_{ij} + \upsilon_i + \zeta_t + \kappa_c + \tau_{ctry} + \varepsilon_{ijct} \qquad (3)$$

where $comp_{ij}$ indicates the competitive relationship between the focal firm i and the reference entity j. If both i and j are focal firms, their technological distance is taken only once to avoid duplications. κ_c is a set of cluster fixed effects, and ζ_t , υ_i and τ_{ctry} are the same dummy variables as defined in equation (2).

Models (6) to (10) shows the results of estimating equation (3) using OLS. For models (6) through (8), the omitted dummy group for $comp_{ij}$ corresponds to dyads where entity j is a nonprofit organization. The results are similar to those obtained by using the average distance: technological distance increases when the reference firm is in the same industry, same segment or direct competition on the product market. In model (9), we explore the distance effect among different types of nonprofit organizations, by changing the omitted dummy group to for-profit firms that are not direct competitors. The results show that technological distance increases when the reference entity is a government agency, which may indicate the specific defense-related research by these institutions. The distance effect is not significant when the reference entity is a university or other nonprofit organization.

The reference organization's local R&D effort is considered in models (10). The dummy variable *comp_core_cluster* is equal to 1 if the competitor conducts the largest percentage of innovations in the cluster. As discussed in Section 2, we expect to see larger technological distance when a focal firm innovates in a competitor's core technological cluster. The positive and statistically significant coefficient on the interaction term *comp_core_cluster* × *competitor* suggests that firms take even more cautious approaches when locating next to the central lab of a competitor, presumably because of the latter's larger absorptive capacity. Co-location in a competitor's core cluster is expected to increase the technological distance by approximately 0.30, twice the increase caused by co-location in a non-core cluster.

4.2 Analysis of Internalization

Table 3 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 dependent variable produce very consistent results.

The total number of innovators in the cluster does not seem to have any significant impact on internalization, even if we only consider for-profit innovators. The effect of competition starts to emerge when we distinguish large from small for-profit innovators. An increase in the number of neighboring large firms increases the self-citation ratio while the opposite is true for small firms. The effect of local competition is more evident when competition is defined in the product market rather than in the technology space. 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 it is that firms self-cite their own patents developed from there. To the extent that self-citations proxy for internalized value, this finding supports our argument that in highly competitive environments, firms tend to develop technologies more integrated with their internal resources.

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 knowledge seeking. 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 is 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 to 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. While there are significant differences in self-citations across quartiles – more self-citations are found in clusters with more competitors – 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 the local environments. Innovation produced in clusters with a strong presence of direct competitors is more tightly intertwined with the firm's internal knowledge base.

4.3 Analysis of Control

Table 4 shows the regression results of *cross_cluster* with negative binomial models. Models (1) to (5) use the total number of local patents as exposure variable, so we essentially test the percentage of local patents that are 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 tendency of using 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 headquarters. Note that we add a new control variable, *core_cluster*, to indicate whether the cluster is the main R&D site for the firm. One would expect that a centralized system exerts more control over geographically dispersed innovation by tightly connecting them with a core R&D center. As expected, the coefficient of *core_cluster* is positive and highly significant.

We further explore cross-cluster links between core and peripheral locations. Models (6) to (10) show the results with a new dependent variable, $to\_cluster_{ict}$, which counts the number of patents in the non-core cluster *c* that have at least one inventor located in firm *i*'s core cluster. The exposure variable used for this estimation is the number of all firm *i*'s cross-cluster patents in cluster *c* ($cross\_cluster_{ict}$); hence, we essentially explore the percentage of cross-cluster patents that are linked to the core 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 that connect to the core cluster, and this effect is stronger when the competitive environment is measured by product-market competition. Therefore, not only does local competition increase the occurrence of cross-cluster links, but the increase is particularly due to connections with the firms' home bases.

4.4 Robustness Checks

The above findings are consistent with our hypothesis that R&D projects in competitive clusters are technologically further away from those of direct competitors, utilized more internally, and 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 include all locations. Cluster membership is determined by the number of desired clusters that we pick region by region to accommodate the 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, 2006). 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 3.

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. Nevertheless, for most models, 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 by the cluster dummies, the results strengthen our belief that firms adjust their local R&D strategies according to the particular competitive environment they are facing.

5. Conclusion

While geographic agglomeration has obvious benefits for firm innovation, it can also bring serious drawbacks. We are interested in exploring how firms are able to tap into the rich resources in technology clusters while still appropriating value from innovation. Our empirical findings suggest that leading firms strategically organize their R&D activities when facing local competitors. A multi-location firm may allocate less vulnerable projects to clustered areas, incorporate local innovations quickly into its global knowledge base, and use cross-cluster teams to intensify the control over locally developed technologies, hence reducing imitation risks. Moreover, firms' strategic responses vary depending on the characteristics of nearby organizations. We find strong evidence of strategic behavior when the neighboring firms share the same product market, but not when they overlap in the technological space.

For firms making location decisions, this study shows that the highly competitive technology clusters are not the forbidden land for industry leaders. Admittedly, the intensive local information flow poses a potential threat to their technological leadership and may even erode their competitive advantage. However, large multinational firms can take advantage of their geographically dispersed organizations to plan their local R&D. The risk of exposing certain technologies to local competitors is low if the local competitors do not have the capabilities to absorb or appropriate these technologies. To take it one step further, because potential knowledge spillovers from the industry leaders tend to attract small firms to cluster around them, avoiding technology clusters is hardly an option for the most technologically advanced. Strategic organization of R&D activities is crucial in such circumstances.

Policy makers who are eager to nurture local high-tech industries often use various incentives, such as tax breaks, to attract firms to conduct R&D there. However, the 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 spillover to the local community. It would be interesting for future research to examine the features of local environments that facilitate not only R&D investments, but also active learning across firm boundaries.

This study also points to several avenues for further inquiries. First, although the mechanisms explored in this paper are based on multi-unit firms, the need to appropriate economic rents from proprietary innovation is universal and 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 appropriate knowledge from technology clusters.

Third, our arguments evolve 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 as well as 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 not only necessary, but also desirable. Moreover, firms' abilities to allocate resources and exercise strategic internalization are limited once innovation goes beyond the same hierarchical structure. Therefore, we need to better understand how firms appropriate value from innovations with permeable, changing, and diffuse firm boundaries.

References

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

Alcácer, J., and Gittelman, M. 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *Review of Economics and Statistics*, forthcoming.

Almeida, P. and Kogut, B. 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45: 905–917.

Anand, B.N. and Galetovic, A., 2004. How market smarts can protect property rights. *Harvard Business Review*, December 2004.

Arora, A., Fosfuri, A. and Gambardella, A. 2001. *Markets for Technology: The Economics of Innovation and Corporate Strategy*. The MIT Press.

Audretsch, D.B. and Feldman, M.B. 1996. R&D spillovers and the geography of innovation and production. *American Economic Review*, 86, 630–640.

Bartlett C.A. and Ghoshal, S. 1990. The multinational corporation as an interorganizational network. *Academy of Management Review*, 15, 603–625.

Baum, J.A.C. and Haveman, H.A. 1997. Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898–1990. *Administrative Science Quarterly*, 42(2): 304–338.

Baum, J.A.C. and Sorenson, O. 2003. Geography and strategy. *Advances in Strategic Management*, Volume 20, JAI Press.

Buckley, P.J. and Casson, C.M. 1976. *The Future of the Multinational Enterprises*. Palgrave-Macmillan.

Buenstrof, G. and Klepper, S, 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. and Veugelers, R. 2002. R&D cooperation and spillovers: Some empirical evidence from Belgium. *American Economic Review*, 92(4): 1169–1184.

Chung, W., and Alcácer, J. 2002. Knowledge seeking and location choice of foreign direct investment in the United States. *Management Science*, 48: 1534–1554.

Cohen, W. and Levinthal, D. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 24: 128–152

Cohen, W. 1995. Empirical studies of innovative activity. In Paul Stoneman (Ed): *Handbook of the Economics of Innovations and Technological Change*. Blackwell Publishers.

Cummings, J.N. 2004. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science*, 50(3): 352–364.

Dierickx, I. and Cool, K.. 1989. Asset stock accumulation and sustainability of competitive advantage. *Management Science*, 35 (12): 1504–1511.

Edstrom, A. and Galbraith, J.R. 1977. Transfer of managers as a coordination and control strategy in multinational organizations. *Administrative Science Quarterly*, 22(2): 248-263.

Feinberg, S. and Gupta, A. 2006. MNC subsidiaries and country risk: Internalization as a safeguard against weak external institutions. Working paper.

Fleming, L., King, C. III and Juda, A. 2006. Small worlds and regional innovation. *Organization Science*, forthcoming.

Gilson, R.J. 1999. The legal infrastructure of high technology districts: Silicon Valley, Route 128, and covenants not to compete. *New York University Law Review*, 74: 575–629.

Hall, B.H., Jaffe, A.B. and Trajtenberg, M. 2005. Market value and patent citations. *Rand Journal of Economics*, 36(1): 16–38.

Henisz, W, and Macher, J. 2004. Firm- and country-level tradeoffs and contingencies in the evaluation of foreign investment: The semiconductor industry, 1994–2002. *Organization Science*, 15(6): 537–554.

Jacobs, J. 1969. *The Economy of Cities*. New York: Random House

Jaffe, A.B. 1986. Technological opportunity and spillovers of R and D: evidence from firms' patents, profits, and market value. *American Economic Review*, 76(5): 984–1001.

Jaffe, A.B., Trajtenberg, M. and Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*. 434, 578–598.

Jaffe, A.B., Trajtenberg, M. and Fogarty, M.S. 2000. The meaning of patent citations: report on the NBER/Case-Western Reserve Survey of patentees. *NBER Working Paper*, No. 7631.

Kogut, B. and Zander, U. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, 3(3): 383–397.

Kogut, B. and Zander, U. 1993. Knowledge of the firm and the evolutionary theory of the multinational corporation. *Journal of International Business Studies*, 24: 625–645.

Gittelman, M and Kogut, B 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citations patterns. *Management Science*, 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 Journal*, 17(Winter Special Issue): 93–109.

Levin, R.C., Klevorick, A.K., Nelson R.R. and Winter, S.G. 1987. Appropriating the returns from industrial R&D. *Brookings Papers on Economic Activity*, 3: 783–820.

Mansfield, E., Schwartz, M. and Wagner, S. 1981. Imitation costs and patents: An empirical study. *Economic Journal*, 91: 907–918.

Marshall, A. 1922. *Principles of Economics*, 8th edition. London: MacMillan.

Nobel, R. and Birkinshaw, J. 1998. Innovation in multinational corporations: control and communication patterns in international R&D operations. *Strategic Management Journal*, 19(5): 479–496.

Oxley, J.E. 1999. Institutional environment and the mechanisms of governance: The impact of intellectual property protection on the structure of inter-firm alliances. *Journal of Economic Behavior and Organization*, 38(3): 283–309.

Oxley, J.E. and Sampson, R.C. 2004. The scope and governance of international R&D alliances. *Strategic Management Journal*, 25(8–9): 723–749.

Porter, M.E. 1985. *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press.

Porter, M.E. 2000. Location, competition and economic development: local clusters in a global economy. *Economic Development Quarterly*, 14: 15–34.

Rajan, R. and Zingales, L. 2001. The firm as a dedicated hierarchy: A theory of the origins and growth of firms. *Quarterly Journal of Economics*, 116 (3): 805–851.

Sanna-Randaccio, F. and Veugelers, R. 2002. Multinational knowledge spillovers with centralized versus decentralized R&D: A game theoretic approach. *CEPR Discussion Paper* 3151.

Saxenian, A. 1996. *Regional Advantage*. Cambridge: Harvard University Press.

Shaver, J.M. and Flyer, F. 2000. Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management Journal*, 21: 1175–1193.

Singh, J. 2005. Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*, 51(5): 756–770.

Swaminathan, A. and Delacroix, J. 1991. Differentiation within an organizational population: Additional evidence from the wine industry. *Academy of Management Journal*, 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 Journal*, 21(8): 791–811.

Stuart, T.E. and Podolny, J. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal*, 17: 21–38.

Stuart, T.E., and Sorenson, O. 2003. Liquidity events and the geographic distribution of entrepreneurial activity. *Administrative Science Quarterly*, 48: 175–201.

Szulanski, G. 1996. Exploring internal stickiness: impediments to the transfer of best practices within the firm. *Strategic Management Journal*, 17: 27–43.

Teece, D., 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy*, 15(6): 285–305.

Trajtenberg, M., Henderson, R., and Jaffe, A. 1997. University versus corporate patents: a window on the basicness of invention. Economics of Innovation and New Technology, 5(1): 19–50. Also in: Jaffe, A. and Trajtenberg, M. (Eds). 2002. *Patents, citations, and innovations*. Cambridge: MIT Press.

Vickers, J. 1986. The evolution of market structure when there is a sequence of innovations. *Journal of Industrial Economics*, 35(1): 1–12.

Yoffie, D.B. 1993. Foreign direct investment in semiconductors. In: K.A. Froot (Ed), *Foreign Direct Investment*. Chicago: University of Chicago Press.

Zhao, M. 2006. Conducting R&D in countries with weak intellectual property rights protection. *Management Science*, 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 Science*, 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
technological_distance	935	0.32	0.17	0	1
cross_cluster	1,089	4.42	9.23	0	90
to_core	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 prduct 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: OLS estimates on technological distance between focal firms and local innovators

Dependent Variable	Average distance					Dyadic distance				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
plants_in_cluster	0.000	0.000	0.000	0.000	0.000					
	(0.15)	(0.21)	(0.64)	(0.25)	(0.41)					
fabless_in_cluster	0.000	0.000	-0.001	-0.001	0.000					
	(1.27)	(1.29)	(1.98)*	(1.75)†	(1.39)					
publications_in_cluster	0.000	0.000	0.000	0.000	0.000					
	(0.49)	(0.21)	(0.43)	(0.39)	(0.30)					
with_plant	0.114	0.114	0.114	0.112	0.112					
	(8.93)**	(8.92)**	(8.98)**	(8.81)**	(8.81)**					
with_fabless	0.027	0.034	0.032	0.018	0.038					
	(0.53)	(0.67)	(0.64)	(0.36)	(0.77)					
innovators_profit	0.001									
	(2.10)*									
small_innovators		0.000								
		(0.23)								
large_innovators		0.002								
		(2.31)*								
in_industry			0.006			0.113				
			(3.68)**			(16.22)**				
not_in_industry			0.000			-0.027				
			(0.19)			(4.53)**				
in_segment				0.003			0.109			
				(2.64)**			(15.64)**			
not_in_segment				0.000			-0.025			
				(0.49)			(4.13)**			
competitors					0.007			0.126	0.142	0.127
					(4.19)**			(15.95)**	(23.83)**	(20.78)**
not_competitors					0.000			-0.015		
					(0.92)			(2.59)**		
innovators_nonprofit	-0.009		-0.007	-0.007	-0.007					
	(2.81)**		(2.20)*	(2.29)*	(2.43)*					
universities		-0.008							0.006	0.006
		(2.35)*							(0.78)	(0.86)
govt_innovators		-0.004							0.051	0.052
		(0.70)							(4.69)**	(4.74)**
other_nonprofit		-0.025							-0.037	-0.037
		(1.78)†							(1.61)	(1.61)
comp_core_cluster										0.011
										(2.82)**
comp_core_cluster × competitor										0.175
										(11.82)**
Constant	0.162	0.163	0.160	0.164	0.165	0.333	0.328	0.345	0.330	0.337
	(2.74)**	(2.75)**	(2.72)**	(2.77)**	(2.81)**	(16.34)**	(16.02)**	(16.77)**	(16.40)**	(17.41)**
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
Cluster fixed effects	N	N	N	N	N	Y	Y	Y	Y	Y
Observations	917	917	917	917	917	21,483	21,483	21,483	21,483	22,218
R-squared	0.21	0.21	0.22	0.22	0.22	0.14	0.13	0.12	0.12	0.14

Absolute value of z statistics in
† significant at 10%; * significant at 5%; ** significant at 1%

Table 3: 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	-0.006	-0.014	-0.010	-0.006	-0.007
	(1.71)†	(1.74)†	(3.73)**	(2.75)**	(1.89)†	(2.12)*
fabless_in_cluster	-0.001	-0.001	-0.002	-0.003	-0.001	-0.001
	(0.77)	(0.85)	(1.90)†	(2.07)*	(0.94)	(0.85)
publications_in_cluster	0.000	0.000	-0.001	-0.000	0.000	0.000
	(0.58)	(0.68)	(1.13)	(0.18)	(0.77)	(0.27)
with_plant	0.132	0.130	0.185	0.169	0.121	0.134
	(2.99)**	(2.92)**	(3.97)**	(3.58)**	(2.70)**	(2.99)**
with_fabless	-0.053	-0.051	-0.159	-0.111	-0.159	-0.077
	(0.19)	(0.18)	(0.58)	(0.40)	(0.56)	(0.28)
patents_total	0.000	0.000	0.000	0.000	0.000	0.000
	(0.18)	(0.17)	(0.68)	(0.19)	(0.25)	(0.15)
patents _semi	0.002	0.002	0.002	0.002	0.002	0.002
	(7.71)**	(7.70)**	(8.13)**	(7.73)**	(7.82)**	(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		-0.003	-0.001	0.002
		(0.25)		(0.26)	(0.10)	(0.16)
universities			-0.009			
			(0.76)			
govt_innovators			0.006			
			(0.18)			
other_nonprofit			0.103			
			(1.40)			
Constant	-17.434	-16.192	-16.903	-17.420	-17.404	-16.167
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(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 4. Negative Binomial estimates on cross-cluster links

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