Distributed R&D, Cross-Regional Knowledge Integration And Quality of Innovative Output

Jasjit Singh
INSEAD, 1 Ayer Rajah Avenue, 138676 Singapore
E-mail: jasjit.singh@insead.edu
Web: http://faculty.insead.edu/singhj
Tel.: +65 6799 5341
Fax: +65 6799 5399

Forthcoming, *Research Policy*

An earlier version of this article won the Haynes Best Paper Prize at the 2005 meetings of Academy of International Business (AIB), and also the Best Paper Award at the 2006 summer conference organized by Danish Research Unit for Industrial Dynamics (DRUID). I thank Ron Adner, Juan Alcacer, Matthew Bidwell, Marco Ceccagnoli, Pushan Dutt, Lee Fleming, Jeff Furman, Javier Gimeno, Phillippe Haspeslagh, Bruce Kogut, Gwendolyn Lee, Aija Leiponen, Josh Lerner, Kwanghui Lim, Ishtiaq P. Mahmood, Jan Rivkin, Myles Shaver, Olav Sorenson, Toby Stuart, Gabriel Szulanski, Stefan Thomke, Ivo Zander, Minyuan Zhao, three anonymous referees and participants in the Singapore Strategy Seminar Series, AIB annual meetings, DRUID summer conference and AOM annual meetings for comments. I also thank Peter Thompson, Nick Argyres and Brian Silverman for sharing their data with me. Finally, I am grateful to INSEAD for funding this research. Errors remain my own.

Abstract

We explore the impact of geographic dispersion of a firm's R&D activities on the quality of its innovative output. Using data on over half a million patents from 1,127 firms, we find that having geographically distributed R&D per se does not improve the quality of a firm's innovations. In fact, distributed R&D appears to be negatively associated with average value of innovations. This suggests that potential gains from access to diverse ideas and expertise from different locations are, on average, offset by difficulty in achieving integration of knowledge across multiple locations. To investigate whether the innovating teams that do manage cross-fertilization of ideas from different locations achieve more valuable innovations, we analyze innovations for which there is evidence of such knowledge cross-fertilization along any of the followings dimensions: knowledge sourcing from other locations within the firm, having at least one inventor with cross-regional ties, and having at least one inventor that has recently moved from another region. Analysis along all three dimensions consistently reveals a direct positive effect cross-regional knowledge integration has on innovation quality, as well as a positive interaction effect of cross-regional knowledge integration and distributed R&D for innovation quality. More generally, our findings provide new evidence regarding the importance of cross-unit integrative mechanisms for achieving superior performance in multi-unit firms.

Keywords: Distributed R&D, Knowledge Integration, Multinational Firms, Collaboration, Mobility

1. INTRODUCTION

Despite the hype about the world becoming more "global", geography continues to play a crucial role in the organization of economic activity (Clark et al, 2000; Sorenson and Baum, 2003). Studies in economic geography have established that acquisition of knowledge from geographically distant sources remains extremely difficult, limiting a firm's ability to build upon non-local knowledge for further innovation (Jaffe et al, 1993; Audretsch and Feldman, 1996; Thompson and Fox-Kean, 2004). However, international business literature suggests that this constraint might be less severe for multi-location firms, since such firms can operate as a distributed innovation network that generates, assimilates and integrates knowledge on a worldwide basis. While geographically dispersed subsidiaries can tap into specialized clusters of expertise dispersed worldwide, formal and informal intra-firm mechanisms can help achieve cross-regional integration in order to make the dispersed knowledge available throughout the firm (Bartlett and Ghoshal, 1989; Hedlund, 1994; Kogut and Zander, 1993; Birkinshaw, 1997; Nobel and Birkinshaw, 1998). However, the empirical question of whether such a mechanism does improve a firm's overall innovative ability still remains largely unexplored. Our present study is an attempt to fill this gap.

In recent decades, firms have placed increasing focus on building a global innovation network, comprised of R&D units located in clusters of excellence around the world (Florida, 1997; Kuemmerle, 1999; Frost et al, 2002; Feinberg and Gupta, 2003). While several empirical studies have established that individual subsidiaries of multi-unit firms do successfully tap into external knowledge sources in their respective locations (Almeida, 1996; Frost, 2001; Singh, 2007), the empirical question of whether this actually improves the firm's overall innovative ability still remains understudied.[1] Only recently have researchers started to study the link between geographic configuration of R&D and innovation performance. In a cross-sectional comparison of 65 Japanese pharmaceutical firms, Penner-Hahn and Shaver (2005) find that firms doing international R&D tend to produce more patents than do firms with only domestic R&D. Furman et al (2006) go beyond this binary choice of whether or not to do

---

[1] Research by Hitt, Hoskisson and co-authors (e.g., Hitt et al, 1997) explores the impact of a firm's international diversification on resources devoted to R&D, but does not examine the impact on actual innovation output.

international R&D by examining the number of R&D locations per therapeutic class that 9 major U.S. pharmaceutical firms have. They find that, holding constant a firm's R&D in a specific therapeutic class, greater number of R&D locations in that class is negatively associated with patent counts. Leiponen and Helfat (2006) analyze survey data on dispersion of R&D by Finnish firms, and similarly find that benefits from dispersed R&D do not apply to novel innovation but only apply to imitative innovation.

While insightful, the above studies do not conclusively establish the link between distributed R&D and performance and make only limited progress regarding cross-regional mechanisms to improve innovative performance of multi-unit firms. Our study extends this literature by not only carrying out a more thorough examination of performance implications of geographic distribution of R&D but also by rigorously examining the role cross-regional knowledge integration plays in the effect that distributed R&D has on innovation. Our study also improves upon previously used methodology in three specific dimensions. First, recognizing large cross-patent variance in quality as well as large cross-firm variance in use of patents to protect intellectual property, we use quality of patents rather than raw patent counts to capture innovative performance. Second, we use longitudinal data analysis to employ appropriate patent-level regression models with firm fixed effects, addressing concerns that unobserved firm heterogeneity could affect results in just cross-sectional models that are commonly employed. Third, by analyzing a sample of over half a million patents from 1,127 major firms across all manufacturing sectors, we can be confident that our findings are not idiosyncratic to a specific industry or a small sample of firms.

Out research complements a recent study by Argyres and Silverman (2004), who investigate the effect of formal R&D organization and control systems on innovative performance of 71 Fortune 500 firms. They find decentralization of formal R&D organization to be negatively associated with average value of patents. However, while they focus on the *formal R&D organization* as the explanatory variable of interest, our focus in this paper is on the *geographic configuration of R&D*. The two issues, though related, are distinct. For example, a firm could have a decentralized formal organization even with relatively small number of R&D locations, while another firm might have a much more centralized organization despite having a much greater number of R&D locations. Further, recognizing the limitation of just looking at formal organizational structure their study focuses on, Argyres and Silverman (2004) call upon future research to also explore informal knowledge integration mechanisms. They suggest that "more nuanced management instruments – such as rotation of personnel, or cross-divisional teams – may be particularly useful in managing such [multi-unit] organizations" (p. 954). Our study takes up this challenge of investigating the link between cross-unit integration and innovative performance of firms.

In studying knowledge integration, we employ directly observed indicators of cross-regional knowledge application within the firm as well as examine specific mechanisms such as personnel rotation and cross-regional interpersonal ties that firms may use to promote such knowledge integration. Thus, our paper is also related to empirical literature on mechanisms promoting knowledge transfer across geographic boundaries. For example, Hansen (1999) and Frost and Zhou (2005) demonstrate that cross-regional interpersonal ties can be an important source of better knowledge flow within a firm. Analogously, Almeida and Kogut (1999) show mobility of individuals as a driver of knowledge diffusion across locations. While this stream of research has established that informal mechanisms such as networks and mobility are indeed important for knowledge flow, it has not examined whether this knowledge flow ultimately translates into more valuable innovations.

Our empirical analysis is based on patents filed by 1,127 firms during the period 1986-1995. Following previous research, we interpret forward citations received by a patent as an indicator of technological and economic value of the underlying innovation. Our findings suggest that, on an average, geographic dispersion of R&D activities does not translate into more valuable innovations. In fact, dispersion of R&D is associated with a non-trivial *decrease* in the average value of a firm's innovations. We explain this result by noting that access to worldwide pockets of knowledge should materialize not from multi-location presence per se but from the firm's ability to actually integrate knowledge dispersed across different locations. In the absence of such integrative mechanisms, potential gains from distributed R&D can easily be overshadowed by increased cost of coordinating multiple units and difficulty of leveraging knowledge from multiple locations. Our empirical analysis reveals that innovations resulting from cross-regional integration of knowledge are indeed of greater value, and that there is also a positive interaction effect of having distributed R&D and being able to achieve cross-regional knowledge integration on value of innovations. These results are robust to using any of three alternate measures of knowledge integration – knowledge sourcing from remote R&D units, cross-regional interpersonal linkages of inventors and mobility of inventors from other locations.

2. DISTRIBUTED R&D, KNOWLEDGE INTEGRATION AND INNOVATION

A knowledge-based view of the firm emphasizes that a firm's accumulated knowledge is key to its continued ability to innovate, and ultimately to its ability to compete (Nelson and Winter, 1982; Grant, 1996; Dosi et al, 2000). Scholars of innovation have argued that novel innovations often result from combination of existing pieces of knowledge (Schumpeter, 1934; Tushman and Rosenkopf, 1992; Nonaka, 1994; Utterback, 1994; Hargadon and Sutton, 1997; Fleming et al, 2007b). Therefore, having a diverse knowledge base within the firm can facilitate innovation through novel combinations of readily accessible pieces of knowledge (Arora and Gambardella, 1990; Patel and Pavitt, 1997; Cohen and Malerba, 2001; Ahuja and Lampert, 2001). As Håkanson (2006) notes: "firms themselves may form epistemic communities in their own right, conferring on their members the means by which specialist knowledge can be effectively combined and integrated" (p. 19).

Novel innovations arise not just from combining ideas from the firm's core areas of expertise but through their recombination with ideas from outside (March, 1991). Thus an effective strategy for innovation needs to overcome the constraint of local search, and to balance exploitation of existing knowledge with non-local exploration for new knowledge (Levinthal and March, 1993; Teece et al, 1997). In this spirit, Tom McKillop (CEO of Astra Zeneca) has remarked "Ninety-nine percent of everything exciting that happens will happen outside your own research labs". Existing literature on exploration has primarily focused on two dimensions of non-local search: search across technological domains and search across firm boundaries (Mowery et al, 1996; Stuart and Podolny, 1996; Rosenkopf and Nerkar, 2001; Rosenkopf and Almeida, 2003). However, another potentially important dimension of search – that across geographical boundaries – remains understudied. There are at least two reasons why geographic search is worth a closer examination. On the one hand, the nature of local expertise that different geographic regions develop over time is often quite distinct and complementary (Porter, 1990; Nelson, 1993; Cantwell and Janne, 1999; Clark et al, 2000; Mahmood and Singh, 2003). On the other hand, it is hard for a firm to readily tap into non-local knowledge sources, since knowledge spillovers tend to be geographically localized (Jaffe et al, 1993; Audretsch and Feldman, 1996; Thompson and Fox-Kean, 2004; Singh, 2005). These observations suggest that, while search across regions can be potentially fruitful for further innovation, it is also quite hard to achieve in practice.

One plausible solution to the above dilemma comes from literature on international business, which attributes the very existence of multinational corporations (MNCs) to their ability to transfer knowledge more effectively than is possible through market-mediated channels (Hymer, 1976; Buckley and Casson, 1976). Building upon this view, a stream of recent literature views the MNC as a geographically distributed innovation network that creates, assimilates and integrates knowledge on a worldwide basis (Bartlett and Ghoshal, 1989; Kogut and Zander, 1993; Hedlund, 1994; Birkinshaw, 1997; Frost et al, 2002). This suggests that having subsidiaries in multiple locations could be an efficient way for conducting non-local search in the geographic space.

The multi-unit firm might be a vehicle that overcomes geographic constraint on knowledge transfer, and promotes worldwide integration of knowledge by combining intra-firm mechanisms for efficient long-distance knowledge transfer with localized knowledge spillovers with other firms and organizations located in different parts of the world. By being simultaneously embedded in knowledge networks in multiple locations, a multi-unit firm can absorb local knowledge from each location better than would have been possible remotely. However, since the ability to absorb such external knowledge depends on the recipient’s own existing cumulative R&D experience (Cohen and Levinthal, 1990), the subsidiaries need to be engaged in R&D activities to effectively absorb knowledge from their respective locations. In other words, getting access to knowledge dispersed globally requires geographically distributed R&D, and not just dispersed subsidiaries with no R&D of their own. While dispersed R&D also risks leakage of a firm’s own knowledge to local players in different locations, large firms are able to restrict such leakage (Singh, 2007; Furman et al, 2006). Mechanisms they use for doing so include colocating far from direct competitors, allocating relatively less vulnerable projects to high-risk clusters and relying on the firm’s worldwide complementary capabilities to leverage knowledge (Shaver and Flyer, 2000; Chung and Alcacer, 2002; Zhao, 2006; Alcacer and Zhao, 2006).

In addition to the benefits from localized knowledge spillovers in multiple clusters of knowledge worldwide, distributed R&D also offers other benefits for innovation. For example, having R&D personnel located in different locations avoids “group think” and limited exploration that would result from locating all of them together. As a result, dispersed subsidiaries tend to develop capabilities that are more distinct, offering greater potential for novel combinations (Zander, 1997; Nobel and Birkinshaw, 1998). A firm with multiple R&D locations might also have better information regarding different markets and customers worldwide, which can also be a source of valuable innovations.

Overall, the above arguments suggest that dispersed R&D should lead to better innovation quality. There might, however, be reasons why such expected gains from distributed R&D do not materialize. To start with, having several small units instead of a single large unit gives up potential benefits of scale and scope (Hirschey and Caves, 1981; Audia et al, 2001). The difficulty of coordinating across distributed R&D units might even have a negative effect on a firm’s innovative performance. Further, in response to local market pressures, dispersed R&D units co-located with regional business units might develop a myopic focus on incremental adaptations demanded locally, rather than maintaining the bigger perspective necessary for truly novel innovations with worldwide application.

Perhaps the most important issue, as Grant (1996, p. 380) emphasizes, is that “the critical source of competitive advantage is knowledge integration rather than knowledge itself”. The arguments favoring dispersed R&D take it for granted that a firm is able to integrate knowledge from its dispersed units in order to achieve novel combinations for further innovation. In reality, however, such integration is not easy. While firms are more efficient than market transactions in transfer and integration of knowledge, cross-regional transfer of complex or tacit knowledge is quite hard even within firm boundaries (Teece, 1977; Kogut and Zander, 1993; Szulanski, 1996; Sorenson et al, 2006). Therefore, even if remote R&D units are exposed to specialized knowledge in their respective regions, a firm’s inability to integrate this dispersed knowledge into its overall knowledge base could prevent the firm's overall innovative capability from improving. On the other hand, the firm still incurs the extra overhead of maintaining dispersed R&D units. This suggests that, rather than there being an unambiguous prediction for the net effect of distributed R&D on a firm's innovative ability, the theoretical prediction is ambiguous. In other words, resolving which of the following competing hypotheses holds in reality becomes an empirical question:

Hypothesis 1a: Geographic dispersion of R&D increases the average value of a firm's innovations.

Hypothesis 1b: Geographic dispersion of R&D decreases the average value of a firm's innovations.

The theoretical prediction is less ambiguous for gains from knowledge integration. The previous arguments clearly suggest that the specific innovations that result from greater integration of knowledge across regions should be of better quality than the innovations that do not. In other words, it is not dispersion of a firm's R&D activities per se, but the ability to combine knowledge across regions that is really fruitful for achieving greater quality of innovations. This motivates our second hypothesis:

Hypothesis 2: Innovations resulting from cross-regional sourcing of knowledge are of greater value.

Firms may rely upon a wide range of formal and informal mechanisms for promoting intra-firm transfer and integration of knowledge (Kogut and Zander, 1992; Ghoshal et al, 1994; Gupta and Govindarajan, 2000). In particular, a key mechanism that facilitates internal flow of knowledge is strong interpersonal network ties that span across different units, these ties being particularly useful when knowledge is complex or tacit (Kogut and Zander, 1993; Hansen, 1999; Hansen and Lovas, 2004; Singh, 2005; Sorenson et al, 2006). Consistent with this view, A.G. Lafley (CEO of P&G) has announced "We want P&G to be known as the company that collaborates – inside and out – better than any other company in the world." There are at least three reasons why the cross-regional setting makes fostering such collaborative ties particularly important. First, the relative infrequency of ties across regions makes ties that do exist particularly crucial for non-redundant transfer of ideas across locations, effectively working as knowledge brokers without which the ideas from different regional networks could not have been combined (Burt, 2004). Second, because cross-regional ties integrate knowledge from different regions, the resulting knowledge is more diverse and the networks are more heterogeneous, allowing a much richer possibility of novel combinations (Reagans and Zuckerman, 2001; Cummings, 2004; Rodan and Galunic, 2004). Third, since individuals in different divisions might otherwise have little economic incentive to help each other, interpersonal ties are particularly useful in reducing transaction costs like opportunism that could otherwise hinder cross-unit transfer and integration of knowledge (Williamson, 1985; Teece, 1986). These arguments suggest our next hypothesis:

Hypothesis 3: Innovations resulting from inventors with cross-regional ties are of greater value.

Intra-regional rotation of personnel can be another important mechanism facilitating knowledge integration across locations. Previous research has suggested a link between labor mobility and knowledge spillovers, though the focus has been on inter-firm rather than intra-firm settings. For example, in his seminal work on the public good aspect of knowledge, Arrow (1962) suggests that cross-firm labor mobility is a source of inter-firm knowledge spillovers that the original firm does not completely get compensated for. Similarly, in a survey of founders of fast growing U.S. companies, Bhide (1994) finds that 71% of these entrepreneurs had built upon ideas they encountered through previous employment. Saxenian (1994) ascribes the free flow of ideas in Silicon Valley to a fluid labor market of engineers, an explanation supported by evidence from Almeida and Kogut (1999) that greater intra-regional knowledge flow occurs in regions with greater labor mobility. Subsequent research has demonstrated personnel mobility as an important mechanism through which knowledge diffuses even over long distances (Rosenkopf and Almeida, 2003, Song et al, 2003; Agrawal et al, 2006).

Given the above evidence, we expect cross-regional mobility of personnel to also facilitate knowledge flow even across different geographic units within the same firm. Further, while none of the above studies directly examines the link between mobility and innovative output, improved innovation quality may be the expected outcome from cross-fertilization of knowledge integration that results from mobility. This leads to our next hypothesis:

Hypothesis 4: Innovations resulting from inventors that move across locations are of greater value.

Combining the arguments we presented when discussing Hypothesis 1 with those for Hypotheses 2 through 4, we can also make predictions regarding a moderating effect between geographic spread of a firm's R&D activities and mechanisms for knowledge integration in determining value of innovations. As R&D activity gets more dispersed, a firm has access to potentially more unique bodies of knowledge through its multi-location presence, offering a richer space for recombination of existing ideas. However, greater distances are also accompanied by increased liability of foreignness and greater difficulty for transmitting tacit knowledge, minimizing detrimental effects from both of which requires a closer integration across units. This suggests that potential value of mechanisms supporting knowledge integration should also increase as the R&D activity itself becomes geographically more dispersed. Put differently, geographic dispersion of R&D provides a platform for tapping into diverse sources of knowledge, but this opportunity is best leveraged when there are also complementary mechanisms that integrate knowledge across different locations. Our final hypothesis captures this interaction effect:

Hypothesis 5: The effect of distributed R&D on innovation is moderated by cross-regional knowledge integration. In other words, greater cross-regional knowledge sourcing, greater cross-regional collaborative ties and greater cross-regional mobility of inventors should all have a positive interaction effect with geographic distribution of R&D in determining value of innovations.

3. DATA

Our dataset was constructed by merging data on patents obtained directly from U.S. Patents and Trademarks Office (USPTO) with additional data fields made available in an National Bureau of Economic Research (NBER) database described by Jaffe and Trajtenberg (2002, Chapter 13).

3.1. Matching Patents to Firms

A challenge in using the patent data is that a single firm may appear in the database as multiple assignee names. For example, a patent arising from a German subsidiary of IBM could appear as "IBM", "International Business machines", "IBM Germany" or a name from which the affiliation to IBM is not obvious at all. This could lead to incorrect inference about a firm's R&D output as well as the geographic distribution of its R&D activity. USPTO data has about 175,000 assignees for 1975-1999, making it hard to verify that each assignee is either indeed an independent entity or gets correctly matched to its ultimate parent. To keep the task manageable, we restricted our attention to about 10,000 of the largest patent assignees. To start with, Compustat-based parent firm identifiers from the NBER database were used to match about 4,600 assignees to their parent firms. Stopford's *Directory of Multinationals* was then used to match over 2,800 additional assignees to their parents. Next, about 400 major government-affiliated bodies, 550 research institutes and 450 universities worldwide were identified as non-firm entities. Finally, the ownership of about 1,000 large assignees was determined using *Who Owns Who* directories and the Internet. The above steps together matched the 10,000 assignees we started with to 3,700 or so unique organizations, only around 2,300 of which are firms and hence of potential interest for our study.

3.2. Sample of Patents

Our main analysis is based upon successful patents applied for during 1986-1995. The number of patents arising during this period was around 1.01 million, only 0.83 million of which are owned by assignees rather than individuals. The 3,700 unique organizations mentioned above account for 0.60 million of these, or about 72% of all assigned patents. The remaining 28% are dispersed among numerous assignees, with no practical way of verifying their ultimate owners. These patents were therefore dropped from our sample. However, the dropped patents would anyway have been of little interest to us, since their assignees typically own too few patents to reveal reliable information regarding geographic distribution of R&D. In fact, for patent-based measures to be more reliable, we further restricted our sample to patents from assignees with at least 20 patents in the previous four years. This led to the final sample having about 0.53 million patents from 1,127 firms. [2]

---

[2] While the regression analysis employs this specific sample, the construction of relevant variables itself uses the entire patent database. For example, measures for value of a patent as well as its technological breadth were constructed using all relevant citations. Likewise, creating a past window of 4 years for a firm's patents required looking at the focal firm's patents even from years before the actual sample period of 1986-1995.

To ensure that our results are not driven by peculiarities of the sample chosen, we carried out robustness checks for all analyses reported in this paper using different samples of firms. First, instead of using a cut-off of 20 recent patents in order for a firm to be included in the analysis, we tried alternate cut-off points of 10 and 50. In addition, to ensure that the results are not driven by systematic differences between patents arising in the U.S. versus those arising overseas, we also tried restricting the sample to patents originating from U.S. inventors and owned by one of the COMPUSTAT firms. Our main findings were found to be robust to all these alternate sampling methods.

3.3. Inventor Data

Our measures of cross-regional collaboration and mobility are based upon data on individual patent inventors. In order to devise a method that minimizes the number of errors in identifying when names on two records really represent the same inventor, we devised a procedure that is a variant of similar algorithms used by Singh (2005), Trajtenberg (2005) and Fleming et al (2007a, 2007b). Specifically, we take two records as having the same inventor if and only if the following hold:

1. The first and last names in the two inventor records matched exactly and the middle initials, when present in both records, are also the same.
2. The two records overlap on at least one of the following dimensions: city, zip code, assignee, technological subclasses, patent citations and collaborating inventors.

In robustness checks not reported in the paper, we find the main findings to be almost identical even if we use the alternate name matching algorithm employed by Singh (2005), which relies upon just the first name, last name, middle initial and technological classification when doing the inventor match.

4. CONSTRUCTION OF VARIABLES

4.1. Quality of a Firm’s Innovations

Only a fraction of innovations actually end up as patents, with firms differing quite a lot in their reliance on patents (Levin et al, 1987). In addition, patents are quite heterogeneous in value, with most of them are worth very little (Jaffe and Trajtenberg, 2002). Therefore, rather than just focusing on patent counts, recent literature has tried to measure the economic and technological importance of patents. In particular, number of citations a patent receives has been shown to be correlated with several direct measures of patent value, including the consumer-surplus generated (Trajtenberg, 1990), expert evaluation of patent value (Albert et al, 1991), patent renewal rates (Harhoff et al, 1999) and contribution to a firm's market value (Hall et al, 2005). Therefore citation-based measures of innovation value have been used in several studies of innovation (Ahuja and Lampert, 2001; Rosenkopf and Almeida, 2003; Argyres and Silverman, 2004). Likewise, we also define *value of innovation* as the number of forward citations received by a patent. [3] It is important to recognize that, since patents from different years have different "windows of opportunity" to be cited in our dataset, a direct comparison of patent citations across patents from different years would be inappropriate. To overcome this issue, we follow Jaffe and Trajtenberg (2002, Chapter 13) in including year fixed effects in all regressions, so that systematic cross-year differences arising from this "truncation bias" are taken into account. Similarly, as described below, technology fixed effects help overcome systematic cross-technology differences in citation rates.

4.2. Geographic Distribution of R&D

Obtaining fine-grained geographic data on a firm's R&D labs is impractical, especially for a large sample of firms. Fortunately, we do observe geographic distribution of the final R&D output, since each patent includes information regarding geographic location of its inventors. This information allowed us to construct two different measures of how distributed a firm's R&D is. In order to ensure that our findings are not driven by the peculiarities of either measure, robustness of all analyses is ensured through replication using both measures. [4]

---

[3] This is consistent with USPTO's view: "If a single document is cited in numerous patents, the technology revealed in that document is apparently involved in many developmental efforts. Thus, the number of times a patent document is cited may be a measure of its technological significance" (Office of Technology Assessment and Forecast, Sixth Report, 1976, p. 167). Our citation-based measure includes both self-citations by the owner firm and citations made by others, since both are signals of patent value: while a self-citation signals that the patent may have helped internal innovation, an outside citation suggests the patent to be a potential source of licensing revenue.

[4] Patent data do not allow making a distinction between locations where a firm has a formal R&D lab versus locations where patents arise outside an R&D lab. However, including patents from even outside formal R&D labs might actually be desirable since they also contribute to the firm's cumulative knowledge base. In any case, this is not a big issue here since neither of our measures of distributed R&D is very sensitive to whether we include patents from locations with a trivial extent of patenting (which are likely to be the locations outside formal R&D labs).

We define our first measure, *R&D geographic spread*, as the natural logarithm of one plus the average geographic distance (in miles) between points of geographic origin (defined as the address of the first inventor) for any two recent patents of the firm. Formally, this average distance is calculated as

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d(i, j)$$

where n is total number of patents that the firm has successfully applied for in the recent four years, and $d(i, j)$ is the distance between the geographic origin for patent i and patent j. Calculation of individual distances using spherical geometry requires first mapping geographical location of the first inventor for each patent to a latitude and longitude on the earth's surface. This mapping was done primarily using Geographic Names Information System (GNIS) of the U.S. Geological Survey for U.S. locations, and using Geonet Names Server (GNS) of the National Geospatial Intelligence Agency for non-US locations. Manual inspection and internet sources were used to further correct common spelling errors, missing locations and name variations. In the end, we were able to precisely match 96% of the inventor city records to specific latitudes and longitudes. For the remaining records, we used the average latitude and longitude for all patents arising in the focal state (for U.S. inventors) or country (for non-U.S. inventors) as an approximation for the inventor's precise geographic location.

Our second measure of geographic distribution of R&D, *R&D dispersion index*, is defined as one minus the Herfindahl of geographic concentration of the firm's patents. Formally, this is calculated as

$$1-\sum_{k}\left(\frac{n_k}{n}\right)^2$$

where n is total number of patents that the firm has successfully applied for in the recent four years, and n_k refers to the subset of these arising from the first inventor being in geographic "region" k. For inventors with U.S. addresses, we follow Jaffe et al (1993) and Thompson (2006) in using metropolitan area as the relevant definition of a "region". Using metropolitan areas has the advantage of capturing geographic organization of actual economic activity, unlike state boundaries that might have been drawn historically for administrative reasons. We used a concordance from Thompson (2006) to map U.S. cities in USPTO data to metropolitan areas.[5] While we would have liked an analogous definition of “region” for patents originating in other countries as well, absence of such data led us to use country as a unit of analysis for the remaining patents.[6]

4.3. Cross-Regional Knowledge Integration

Our first measure of knowledge integration uses citations to previous patents as an indicator of knowledge flow (Jaffe and Trajtenberg, 2002). A concern is that the inventor may sometimes use citations simply to avoid litigation or clarify claims. Also, a citation added by patent examiners might not reflect knowledge flow (Alcacer and Gittelman, 2006; Thompson, 2006). Despite these sources of noise, however, citations have been shown to correlate well with actual knowledge flow (Jaffe and Trajtenberg, 2002, Chapter 12; Duguet and MacGarvie, 2005). The above concerns are also less serious for self-citations by a firm to the extent that a firm surely knows about its own previous patents. Therefore, previous research has often used self-citations to capture within-firm knowledge flow (Rosenkopf and Nerkar, 2001; Frost and Zhou, 2005). Likewise, we define a binary variable *cross-regional knowledge sourcing* as being 1 if and only if the focal patent makes a backward citation to a patent originating in another geographic unit of the same firm.

Our second measure of knowledge integration is based on cross-regional interpersonal ties that could promote knowledge flow. One way of inferring such ties is analyzing past joint projects as reflected in research papers or patents (Cockburn and Henderson, 1998; Newman, 2001). Fleming et al (2007a) present field-based evidence that collaborations recorded on patents do capture strong interpersonal ties, since co-inventors of a patent typically have a close relationship that persists long after the joint project. Consistent with this view, many recent papers have found past collaborative ties

---

[5] For U.S. cities that could not be mapped to metropolitan areas, we followed the convention of defining a “phantom area” per state. We also tried using state boundaries rather than metropolitan areas as a robustness check, and the results were very similar.

[6] Patent counts for a typical U.S. metropolitan area are of the same order of magnitude as those for a typical foreign country. For example, of the ten largest “regions” in terms of patent output, five are U.S. metropolitan areas and five are foreign countries. Any concerns that we might be comparing “apples and oranges” in defining regions are further mitigated using a dummy for non-US inventors in all regressions, and also by checking the robustness of the regression finding to using only patents arising from U.S. inventors.

to be a strong predictor of future knowledge flow (Frost and Zhou, 2005; Singh, 2005; Sorenson and Singh, 2007). Likewise, we define a dummy variable *cross-regional tie* to be 1 if and only if at least one of the inventors of the focal patent had collaborated with someone from another region in the four years preceding the current patent.

Our final measure of cross-regional knowledge integration is based on individuals moving across regions. Others have inferred such mobility through changes in an inventor's recorded address in patents filed over time (Almeida and Kogut, 1999; Rosenkopf and Almeida, 2003; Song et al, 2003; Agrawal et al, 2006). To construct an analogous measure for mobility across regions, we define *cross-regional move* as a dummy variable that equals 1 for the focal patent if and only if at least one of its inventors have moved from another region in the last four years.

4.4. Control Variables

The size of a team can have a direct effect on the quality of its innovations because of economies of specialization, a larger and more diverse pool of knowledge, and access to a wider and more heterogeneous external network. Therefore, an important patent-level control variable is *team size*, which measures the number of researchers in the innovating team for the focal patent.

Patents based on a broad range of technologies could potentially draw from a richer universe of combinations. We account for this through a control variable *technological breadth* undertaken for an innovation, defined for the focal patent i as

$$1-\sum_j s_{ij}^2$$

where s_{ij} refers to the fraction of patents cited by patent i that belong to technology class j (Jaffe and Trajtenberg, 2002; Argyres and Silverman, 2004). Since this variable is undefined when there are no backward citations, our regressions deal with missing values by setting *technological breadth* to 0 for these observations and setting a dummy *no backward citations* as 1 to capture this special case.

Our analysis also uses two control variables derived with one year lag from COMPUSTAT: *firm size*, the total number of employees in the focal patent's owner firm, and *R&D intensity*, the ratio of R&D

to sales for the firm. We avoid dropping non-COMPUSTAT firms in the regressions by setting these variables to be 0 and defining corresponding dummy variables *missing firm value* or *missing R&D intensity* respectively as 1. The main results are very similar even if, instead of doing this, we either simply exclude these two variables from the regressions or just include firms with no missing values.

We also construct three control variables related to the firm's patent stock. First, *domain patent stock* is the number of patents a firm accumulated in the given technology in the previous four years, as that might affect its absorptive capacity (Cohen and Levinthal, 1990). Second, we define *worldwide patent stock* as the overall number of patents for the focal firm in the previous four years, as that could either increase innovation quality through more recombination possibilities or greater perceived technological prowess (Podolny et al, 1996) or decrease it by making the firm more inward looking and myopic in its search. Finally, we construct a similar variable *subsidiary patent stock* to allow the scale effect to operate at the level of a subsidiary rather than just the whole firm.

We also control for characteristics of the region where a patent originates. Since patents from larger regions may receive more citations because of localized knowledge flow, we create a variable *region patent stock* as the patent count for a region in the last four years. Second, citation counts for U.S. versus non-U.S. inventors could differ because of systematic differences in patenting strategies of firms, knowledge diffusion patterns, extent of examiner-added citations, and other country-specific factors. To account for this, we define a dummy variable *non-US inventor* as 1 if and only if one of the inventors of a patent resides outside the U.S. Finally, MNCs may strategically manage the nature of their R&D activity in low-IPR countries to prevent others from exploiting their local R&D efforts (Zhao, 2006; Alcacer and Zhao, 2006). Since this could affect citation patterns, we use a variable *IPR protection* that captures the extent of IPR protection in a given country (Ginarte and Park, 1997). This variable is defined as the Ginarte-Park index of IPR protection for the focal patent inventors' country (for the closest available year), with the lowest value chosen in case of patents with inventors based in multiple countries.

5. ANALYSIS AND RESULTS

5.1. Descriptive Statistics

Since our sample extends over multiple years, characteristics of a firm vary within the sample. Nevertheless, to give a sense of the 1,127 firms we analyze, Table 1 uses the median values across all years for each firm as a basis for calculating some overall firm-level summary statistics. The average distance between pairs of recent patents for a firm (calculated using distance between cities for the first inventors for all patent pairs in the last four years) varies between 0 miles and 4,988 miles, with the mean being 638.4 miles. The mean of the Herfindahl measure for geographic concentration of a firm's patents is 0.61. The firms have 12.1 patenting locations on an average, the minimum being 1 and the maximum being 132.[7] The mean value for fraction of patents with the first inventor being outside the U.S. is 0.44. The mean patent stock is 189.7 patents worldwide, and the smallest patent stock is 20 just by the way the sample is constructed (as described earlier).

Table 2 summarizes the definitions and patent-level summary statistics of variables used in the regression analysis below. In analysis by year not reported in the table, we found that the number of R&D locations per firm remained relatively stable during the period we study. However, the extent of cross-regional integration of knowledge grew significantly over time. As Figure 1 illustrates, the fraction of patents for which *cross-regional collaboration* is 1 rose from 19% in 1986 to 35% in 1995. Similarly, the fraction of patents for which *cross-regional move* is 1 rose from 10% in 1986 to 17% in 1995. Finally, the fraction of patents for which *cross-regional knowledge sourcing* is 1 went up slightly from 14% in 1986 to 17% in 1995. Thus, all three measures point to an increase in cross-regional knowledge integration over time. While interesting, these trends are not the focus of this paper. In fact, to ensure that our

---

[7] The number 132 might sound particularly high for the number of locations. However, this arises (in this case, for General Electric) since the statistic only reports the number of unique locations where a firm has one or more patents arising. Given the sensitivity of this statistic to having locations with a trivial extent of patenting, we do not use this directly as a measure of distributed R&D in regression analysis. Both our actual measures of geographic distribution of R&D (*R&D geographic spread* and *R&D dispersion index*) are relatively insensitive to including versus excluding such "outlier" locations that exhibit a trivial extent of patenting.

findings are not driven just by time trends (possibly driven by unobserved factors), we explicitly include year fixed effects in all regression models.

Table 3 gives correlations among the key variables. In the absence of relevant control variables, these correlations give few insights regarding what drives *value of innovations*. In particular, the number of citations received by patents from different years and technologies is not directly comparable, an issue we address in our regression analysis. Also note that many of the explanatory and control variables are quite correlated. While this could have raised questions about multicollinearity, such concerns are mitigated by large sample sizes that make estimation possible without blowing up the standard errors.

5.2. Regression Models

Since the dependent variable *value of innovation* and most explanatory and control variables differ across even contemporary observations within the same firm, we take a patent as the unit of analysis. This prevents loss of information as well as any unexpected biases that may result from aggregation of data. In addition, staying at the patent level allows use of negative binomial regression models, which is desirable given the count data nature of the underlying dependent variable. As in other similar studies, an econometric concern would be that unobserved firm heterogeneity could drive the results. To account for this, we employ negative binomial models with firm fixed effects, which exploit the longitudinal nature of the data to do the estimation using within-firm comparison only.[8]

It is also important to account for differences in observed citation frequency driven by factors other than actual value of the innovation (Jaffe and Trajtenberg, 2002). First, number of citations per patent is incomparable as a measure of quality across technologies. Therefore, all regressions include fixed effects for the focal patent's technological "subcategory" as defined in the NBER data.[9] In addition, patents from later years in our sample may receive fewer citations not necessarily because they are of

---

[8] We also repeated all analyses using pooled negative binomial regressions with robust standard errors, while clustering on firm identifier to avoid any assumption of independence of patents drawn from the same firm. The main findings were qualitatively similar to those reported here.

[9] We also tried using fixed effects for a patent's 3-digit technology primary class instead of 2-digit subcategory. Given that there are more than 400 classes (as opposed to 36 subcategories), doing so in a negative binomial model turned out to be impractical with firm fixed effects. Therefore, as a more practical robustness check, we repeated the analysis using fixed effects for a 3-digit class while not using firm fixed effects, and the main results did not change.

lower quality but simply because of the shorter observed life of these patents. This issue is resolved by using year fixed effects in all regressions.

5.3. Effect of Distributed R&D on Value of Innovation

The first part of our regression analysis, reported in Table 4, examines how the level of dispersion of a firm's R&D activities affects the value of its innovations. While columns (1) and (2) report results based on the distance based measure, *R&D geographic spread*, columns (3) and (4) analyze robustness to using instead the Herfindahl-based measure, *R&D dispersion index*. In all cases, a Wald test confirms that our explanatory variables as well as our control variables add significant explanatory power to the model.

The findings are remarkably consistent across all four columns. Contrary to the prediction of Hypothesis 1a, the effect of distributed R&D on value of a firm's innovation is never positive. In fact, consistent with the competing Hypothesis 1b, the effect is always negative and statistically significant. Since the regression model is negative binomial, the magnitude of this effect needs to be computed separately. For example, the coefficient estimate for *R&D geographic spread* in column 2 implies that increasing the average distance between a firm's recent patent inventors from 0 miles to 1000 miles leads to a 1.2 decrease in the *value of innovation*. Given that the median patent receives 3 citations, this is a significant decrease. As a benchmark for this economic significance, if we consider the finding by Hall et al (2005) that a firm's market value decreases by 3% for a decrease by one in the average number of citations to its patents, the 1.2 decrease reported above would mean a 3.6% decrease in the firm's stock market value. While such a literal interpretation of the estimated negative effect should be treated with caution, the evidence appears strong enough at least to conclude that the effect of distributed R&D per se on the *value of innovation* is definitely not positive!

While the control variables included are not our central focus in this paper, some findings are still worth pointing out. Consistent with our expectation based on previous literature, the *value of innovation* is typically greater for innovations involving larger teams, greater technological breadth, greater domain knowledge, and greater overall regional knowledge stock. The value is typically smaller for innovations with no backward citations, those involving inventors from outside the U.S., and those involving low-IPR countries. Interestingly, once the domain-specific patent stock has been accounted for, a firm's overall patent stock is negatively associated with the *value of innovation*. The most counter-intuitive finding is perhaps the negative association of a firm's R&D intensity on the *value of innovation*.[10] One explanation could be that, though the extent of R&D increases the number of innovations, the average value of these innovations could remain constant or even fall as the marginal R&D dollar gets spent on projects with increasingly lower pay-off.

5.4. Effect of Cross-Regional Knowledge Integration on Value of Innovation

The next part of our analysis examines whether innovations that rely upon cross-regional knowledge integration have a greater *value of innovation*. Columns (1), (2) and (3) in Table 5 report the results for Hypotheses 2, 3 and 4 respectively, considering in turn *cross-regional knowledge sourcing*, *cross-regional tie* and *cross-regional move* as measures for the extent of knowledge integration. In line with these hypotheses, all three dimensions of knowledge integration are found to have a significant positive association with the *value of innovation*. These effects are not just statistically significant but also reflect non-trivial economic magnitude. For example, the implied marginal effects for column (1) suggest that innovations involving a citation to another geographic unit within the firm have a value that is greater than the rest by around 0.5. Since the median patent in our sample receives only 3 citations, this represents a 17% increase in value for the median patent. The results weaken somewhat once we do the same calculation for the other two measures of knowledge integration. Innovations involving at least one inventor with recent cross-regional ties have a value that is about greater than those that do not by about 0.24, which reflects an 8% increase in value. The estimated magnitude of increase in the *value of innovation* is smallest for the measure involving recent movement of at least one inventor from another region, in which case the 0.12 improvement in patent value implies only a 4% expected increase in value.

Even more interesting is the analysis of moderating effect suggested in Hypothesis 5, wherein we should expect distributed R&D to be more beneficial when there is also evidence of cross-regional

---

[10] The same result holds even if we employ ln(R&D expenditure) rather than R&D intensity in the regressions.

knowledge integration. The results are reported in columns (4) through (6) of Table 5, where the *R&D geographic spread* variable is interacted with *cross-regional knowledge sourcing*, *cross-regional ties* and *cross-regional move* respectively as alternate measures of cross-regional knowledge integration. In all three cases, the interaction term is positive and significant. Consistent with Hypothesis 5, gains from distributed R&D are indeed greater when cross-regional knowledge integration is achieved. Interestingly, the magnitude of this interaction effect in itself is never enough to completely overturn the negative effect of distributed R&D, though the overall effect does sometimes turn positive when the direct effect of cross-regional knowledge integration is also taken into account.

To ensure that the findings reported in Table 5 are not driven by any peculiarities of the *R&D geographic spread* measure for distributed R&D, Table 6 uses *R&D dispersion index* as an alternate measure. The main findings regarding both direct and interaction effects continue to hold, with the only exception being that the interaction effect with *cross-regional move* (column 6) is still positive but no longer statistically significant at conventional levels.

6. DISCUSSION AND IMPLICATIONS

Our empirical analysis finds that dispersion of R&D activities, instead of having a positive effect on the quality of innovation, has a statistically significant and economically non-trivial negative effect on the value of innovations a firm generates. This suggests that the extent of cross-regional integration that happens in reality is not, at least on average, enough to overcome the difficulty of managing and integrating knowledge across dispersed R&D units. This interpretation is further boosted by our finding that innovating teams that do achieve cross-regional knowledge integration come up with innovations of significantly greater value than those that do not. The estimated extent of the difference is between 4% and 17%, depending upon which of our three measures of knowledge integration is used for estimation. The 17% estimate based on cross-regional backward citations made by a patent is perhaps most meaningful here, as it directly captures the effect of knowledge sourcing across R&D units on the value of innovation. The other two measures based on cross-regional ties or moves, for which the estimates are lower, only capture conditions under which knowledge integration *could* have occurred.[11] The exact numerical interpretation of these estimates should, however, be treated with caution since patent-based variables can be quite noisy in capturing the underlying constructs for interpersonal ties, knowledge diffusion and innovation quality.

We find cross-regional knowledge integration not only to have a direct effect on quality of innovation but also to have a moderating effect on geographic distribution of R&D for determining innovation quality. The findings suggest that the observed negative association between distributed R&D and innovation quality could be a result of insufficient cross-regional integration for geographically dispersed knowledge to translate into better innovation performance of the firm as a whole. If firms could further improve their use of integrative mechanisms like cross-regional collaborations and personnel rotation, the observed negative average effect of distributed R&D on the value of innovations could disappear and possibly even get reversed. These conclusions echo the findings from a recent survey of 186 major R&D firms, conducted jointly by INSEAD and Booz Allen Hamilton. A report by Doz et al (2006) summarizes the survey's key findings:

> "The knowledge inputs required for innovation in most sectors are increasingly dispersed. Innovation and R&D activities are mirroring this trend with sites being opened or scaled up in locations which will give companies access to differentiated knowledge through accessing internal and external capabilities as well as markets and customers. However, few companies are currently reaping the gains from this globalization: Co-ordination across sites is weak; harmonized processes and systems are lacking; and the need for a cadre of people with international experience is being largely overlooked. There are also significant opportunities for companies to optimize the diversity of their external network by looking further a field for collaborators and to improve their internal knowledge mobility by proactively encouraging internal knowledge sensing and knowledge re-use."

These remarks are consistent with the findings from our study. Like this survey, our analysis also reaches the conclusion that the bottleneck in realizing full potential from access to a global pool of knowledge is not the presence of dispersed R&D sites per se but the mechanisms to integrate knowledge across these R&D sites. There is some evidence that, over time, even firms themselves come to realize the fact that

---

[11] As a check that this is indeed the case, we tried a regression with *cross-regional knowledge sourcing* as dependent variable and *cross-regional tie* as well as *cross-regional move* as independent variables. As expected, the latter two had positive and significant coefficient estimates, but explained only a fraction of the overall variance.

they need to do a better job of cross-unit knowledge integration. That might in part explain why, as pointed out earlier in our discussion of Figure 1, the fraction of innovations resulting from cross-border knowledge integration shows an upward time trend even during our sample period.

The above discussion has two important implications for R&D management. The first imperative is to have incentives and processes in place for employees to avoid reinventing the wheel in innovation, and instead devote more time and energy into sensing knowledge what already exists in the external environment. The second imperative for R&D management is to ensure that the sensing process is not restricted to the geographic location of a specific R&D unit but also includes searching for knowledge scattered in other locations where the firm also has R&D units. As Hansen and Nohria (2004) also report, it appears that firms are becoming sensitive to the first imperative mentioned above, though the second issue is still not being given due attention by most firms. In particular, there is an urgent need for firms to cultivate a larger breed of people that are not only adept in their own domains but also have substantive exposure to and inter-personal ties across different regions. Such individuals can act as bridges to facilitate integration of knowledge across regions.

One constraint we faced in the current study was the inability to simultaneously examine formal R&D organization and informal integrative mechanisms emphasized here. As a robustness check to ensure that our results are not driven by an omitted variable bias arising from not accounting for formal organization of R&D, we merged our data with information on formal R&D organization we obtained from Argyres and Silverman (2004). We were able to match 58 of the 71 firms in their sample with firms in our sample. Although this comprises a small fraction of the overall sample of 1,127 firms we study, it was reassuring that our main findings continued to hold in this sub-sample even after accounting for organization structure variables (*centralized*, *centralized hybrid*, *balanced hybrid*, *decentralized hybrid* and *decentralized*) from Argyres and Silverman (2004). Going forward, we see a promising research opportunity in more closely examining the interaction between formal R&D organization and integrative mechanisms such as cross-unit collaboration and mobility of individuals.

While the specific focus of this paper has been distributed R&D, it also makes an important contribution to research on multi-unit and multi-national organizations in general. While several researchers have tried to examine performance drivers in multi-unit firms, data constraints often prevent such studies from getting at the micro-level mechanisms behind observed performance effects. As Audia et al (2001, p. 99) point out, “The direct examination of factors underlying the multi-unit effect – for example, knowledge transfer across units and bureaucratization - strikes us as an interesting line of investigation. Although an in-depth understanding of how multi-unit firms function offers clear benefits, the difficulty of this approach lies in obtaining direct measures of these underlying processes.” This article takes a step in this direction by using data on cross-border patent citations, inventor collaborations and personnel mobility to directly measure micro-level mechanisms affecting performance. The general point that emerges is that, to really gain from multiple locations and justify the overhead of coordinating multiple units within the same firm, a firm needs to actively manage positive spillovers across individual units. Further, achieving net positive spillovers is often not just a matter of having the right formal organizational structure and controls, but also informal mechanisms that promote cross-regional integration and application of a firm’s knowledge and capabilities. While research in international business has often emphasized the importance of leveraging resources and capabilities distributed among divisions around the globe (Bartlett and Ghoshal, 1989; Birkinshaw, 1997; Nobel and Birkinshaw, 1998; Frost et al, 2002), we demonstrate that expected positive performance effects from distributing activities worldwide might not accrue in the absence of appropriate integrative mechanisms.

7. LIMITATIONS AND FUTURE RESEARCH

While we believe that our study makes an important contribution to understanding performance implications from distributed R&D and cross-regional knowledge integration, it has some limitations. The issues raised in this section should therefore be kept in mind when making normative recommendations based on our findings. More importantly, since these issues are not specific to our study but arise in numerous empirical studies like ours, we discuss them here in the hope that future research will be able to address these in order to push this stream of research even further.

Our regression analysis includes checking for robustness to firm fixed effects, which is important given that purely cross-sectional as well as even pooled regression models can often be plagued by biases resulting from unobserved firm heterogeneity. However, even a fixed effects model does not completely resolve the issue of unobserved *time-varying* characteristics of firms, and the related issue of endogeneity of R&D locations. A firm's decision to invest in certain geographic areas in certain time periods may be a result of time-varying factors that we do not observe as econometricians but that a firm's managers observe and recognize as affecting firm performance. While the time lag employed in constructing explanatory and control variables mitigates such concerns to some extent, it does not completely eliminate them. To the extent that such unobserved factors could be driving observed within-firm differences in quality of innovation, care should be exercised in interpreting the observed relationship between distributed R&D and quality of innovation as being causal.

Our analysis of knowledge integration has some limitations as well. Although we use multiple measures to ensure robustness of our findings, such patent-based are admittedly still somewhat crude and fail to capture all intricacies of knowledge integration one could explore through more in-depth study of individual firms (e.g., Orlikowski, 2002). Data limitations also prevent us from accounting for the possibility that the need for mechanisms for cross-regional knowledge integration might vary with characteristics like tacitness of the underlying knowledge (Subramaniam and Venkatraman, 2001). In addition, since achieving cross-regional knowledge sourcing and maintaining cross-regional ties can be costly, it is plausible that such investments get made only for projects that are particularly promising. Thus an observed association between knowledge integration and quality of innovation need not always be causal, and could instead result from choices managers make regarding which projects to invest in and how to manage them. Likewise, a positive association between cross-regional mobility and quality of innovation need not always reflect a causal relationship, since it could also result from good managers being more likely to be assigned to projects in diverse locations. We hope that future research will try to address these and other related methodological challenges in advancing this stream of research.

REFERENCES

Agrawal, A., I. Cockburn and J. McHale, 2006. Gone But Not Forgotten: Labor Flows, Knowledge Spillovers, and Enduring Social Capital. *Journal of Economic Geography*, Forthcoming.

Ahuja, G. and C.M. Lampert, 2001. Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, 22: 521-543.

Albert, M.B., D. Avery, F. Narin and P. Mcallister, 1991. Direct Validation of Citation Counts as Indicators of Industrially Important Patents. *Research Policy* 20: 251-259.

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

Alcacer, J. and M. Zhao, 2006. Understanding Local Clusters through Global Firms: The Semiconductor Industry. Mimeo.

Almeida P., 1996. Knowledge sourcing by foreign multinationals: Patent citation analysis in the U.S. semiconductor industry. *Strategic Management Journal* 7: 155-165.

Almeida, P. and B. Kogut, 1999. “The Localization of Knowledge and the Mobility of Engineers in Regional Networks,” *Management Science*, Vol. 45(7), 905-917.

Argyres, N.S. and B.S. Silverman, 2004. R&D, organization structure and the development of corporate technological knowledge. *Strategic Management Journal* 25(8-9): 929-958.

Arora, A. and A. Gambardella, 1990. Complementarities and external linkages: the strategies of large firms in biotechnology. *Journal of Industrial Economics*, 38(4): 361-379.

Arrow, K. J., 1962. Economic welfare and the allocation of resources for invention. In R. R. Nelson (ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Vol. 13 of NBER Special Conference Series, Princeton University Press, New Jersey, pp. 609-625.

Audia, P.G., O. Sorenson and J. Hage, 2001. Tradeoffs in the organization of production: multiunit firms, geographic dispersion and organizational learning. In: J. A. C. Baum & H. R. Greve (Eds), *Multiunit Organization and Multimarket Strategy: Advances in Strategic Management* 18:75-105.

Audretsch, D.B. and M.P. Feldman, 1996. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review* 86(3): 630-640.

Bartlett, C.A. and S. Ghoshal, 1989. *Managing Across Borders: The Transnational Solution*. Harvard Business School Press: Boston, MA.

Bhide, A. 1994. How entrepreneurs craft strategies that work. *Harvard Business Review* 72: 150-161.

Birkinshaw, J.M. 1997. Entrepreneurship in multinational corporations: the characteristics of subsidiary initiatives. *Strategic Management Journal* 18(3): 207-229.

Buckley, P.J. and M.C. Casson. 1976. *The Future of the Multinational Enterprise.* London: Holmes & Meier.

Burt, R.S., 2004. Structural Holes and Good Ideas. *American Journal of Sociology* 110(2): 349-399.

Cantwell, J.A. and O. Janne, 1999. Technological globalisation and innovative centres: the role of corporate technological leadership and locational hierarchy. *Research Policy*. 28:119-144.

Chung, W., and J. Alcacer. 2002. Knowledge Seeking and Location Choice of Foreign Direct Investment in the United States. *Management Science* 48(12):1534-1554.

Clark, G.L., M. P. Feldman and M. S. Gertler (Eds.), 2000. *Oxford Handbook of Economic Geography.*

Cockburn, I.M. and R.M. Henderson, 1998. Absorptive Capacity, Coauthoring Behavior, and the organization of Research in Drug Discovery. *Journal of Industrial Economics*, 46(2): 157-182.

Cohen, W.M. and D.A. Levinthal, 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35: 128-152.

Cohen, W.M. and F. Malerba, 2001. Is the tendency to variation a chief cause of progress? *Industrial and Corporate Change* 10: 587-608.

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

Dosi, G., Nelson, R.R. and S.G. Winter, 2000. *The Nature and Dynamics of Organizational Capabilities*. Oxford University Press, Oxford.

Duguet, E. and M. MacGarvie. 2005. How well do patent citations measure knowledge spillovers? Evidence from French innovation surveys. *Economics of Innovation and New Technology* 14(5).

Doz, Y., K. Wilson, S. Vedhoen, T. Goldbrunner and G. Altman, 2006. Innovation: Is Global the Way Forward? Mimeo.

Feinberg, S., and A.K. Gupta, 2003. Knowledge Spillovers and the Assignment of R&D Responsibilities to Foreign Subsidiaries. *Strategic Management Journal*.

Fleming, L., C. King and A. Juda, 2007a. Small Worlds and Innovation. *Organization Science*, Forthcoming.

Fleming, L., S. Mingo and D. Chen, 2007b. Brokerage and Collaborative Creativity. *Administrative Science Quarterly*, Forthcoming.

Florida, R., 1997. The globalization of R & D: results of a survey of foreign-affiliated R&D laboratories in the USA. *Research Policy* 26(1): 85-103.

Frost, T.S., 2001. The geographical sources of foreign subsidiaries' innovations. *Strategic Management Journal* 22: 101-123.

Frost, T.S., J. Birkinshaw and P. Ensign, 2002. Centers of excellence in multinational corporations. *Strategic Management Journal* 23: 997-1018.

Frost, T.S. and C. Zhou, 2005. R&D Co-Practice and 'Reverse' Knowledge Integration in Multinational Firms. *Journal of International Business Studies* 36: 676-687.

Furman, J., M. Kyle, I. Cockburn and R. Henderson, 2006. Public and Private Spillovers, Location & the Productivity of Pharmaceutical Research. *Annales d'Economie et de Statistique,* forthcoming.

Ghoshal, S., H. Korine and G. Szulanski. 1994. Interunit communication in multinational corporations. *Management Science* 40: 96-110.

Ginarte J. and W.G. Park. 1997. Determinants of patent rights: A cross-national study. *Research Policy* 26 (3): 283-301.

Grant, R.M., 1996. Prospering in dynamically-competitive environments: organizational capability as knowledge-integration. *Organizational Science*, 7(4): 375-387.

Gupta, A.K. and V. Govindarajan, 2000. Knowledge flows within multinational corporations. *Strategic Management Journal* 21(4): 473-496.

Håkanson, L. 2006. The Firm as an Epistemic Community: A Research Agenda. Mimeo.

Hansen, M.T., 1999. The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly* 44: 82-111.

Hansen, M.T. and B. Lovas, 2004. How do multinational companies leverage technological competencies? Moving from single to interdependent explanations. *Strategic Management Journal* 25: 801-822.

Hall, B. H. and M. Trajtenberg, 2004. Uncovering General Purpose Technologies with Patent Data. *National Bureau of Economic Research* Working Paper 10901.

Hall, B.H., A.B. Jaffe and M. Trajtenberg, 2005. Market Value and Patent Citations. *Rand Journal of Economics* 36.

Hansen, M. and N. Nohria. 2004. How to build collaborative advantage? *Sloan Management Review* 22-30.

Hargadon, A. and R. Sutton, 1997. Technology Brokering and Innovation in a Product Development Firm. *Administrative Science Quarterly* 42: 716-749.

Harhoff, D., F. Narin, F.M. Scherer and K. Vopel, 1999. Citation Frequency and the Value of Patented Inventions. *The Review of Economics and Statistics* (193): 511-515.

Hedlund, G., 1994. A model of knowledge management and the N-form corporartion. *Strategic Management Journal* 15: 73-90.

Henderson, R., A. Jaffe and M. Trajtenberg, 1998. Universities as a source of commercial technology: a detailed analysis of university patenting, 1965-1988. *Review of Economics and Statistics* 80: 119-127.

Hirschey R and R.E. Caves, 1981. Research and transfer of technology by multinational enterprises. *Oxford Bulletin of Economics and Statistics* 43(2): 115–130.

Hitt, M.A., Hoskisson, R.E. and H. Kim, 1997. International Diversification: Effects on Innovation and Firm performance in Product-Diversified Firms. *Academy of Management Journal* 40(4):767-798.

Hymer, S.H. 1976. *The International Operations of National Firms: A Study of Direct Investment.* MIT Press, Boston, MA.

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

Jaffe, A.B. and M. Trajtenberg, 2002. *Patents, Citations & Innovations: A window on the knowledge economy*. MIT Press, Cambridge.

Kogut, B. and U. Zander, 1992. Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science* 3 (3): 383-397.

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

Kuemmerle, W., 1999. Foreign direct investment in industrial research in the pharmaceutical and electronics industries – results from a survey of multinational firms. *Research Policy*.

Leiponen and Helfat, 2006. Geographic location and decentralization of innovative activity. Mimeo.

Levin, R., A. Klevorick, R. Nelson and S. Winter, 1987. Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity* 3:783-820.

Levinthal, D.A. and J.G. March, 1993. The myopia of learning. Strategic Management Journal 14: 95-112.

Mahmood, I.P. and J. Singh. 2003. Technological Dynamism in Asia. *Research Policy* 32 (6), pp. 1031-1054

March, J.G., 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2(1): 71-87.

Mowery, D.C., J.E. Oxley and B.S. Silverman, 1996. Strategic Alliances and Inter-firm Knowledge Transfer. *Strategic Management Journal* 17: 77-91.

Nelson, R. and S. Winter, 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge.

Nelson, R.R. (ed.), 1993. *National Innovation Systems*. Oxford University Press: New York.

Newman, M.E.J., 2001. The Structure of Scientific Collaboration Networks. Proceedings of U.S. *National Academy of Science* 98: 404-409.

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

Nonaka, I., 1994. Dynamic theory of organizational knowledge creation. *Organization Science* 5(1): 14-37.

Orlikowski, W.J., 2002. Knowing in Practice: Enacting a Collective Capability in Distributed Organizing. *Organization Science*, 13: 249 - 273.

Patel, P. and K. Pavitt, 1997. The technological competencies of the world's largest firms: complex and path dependent, but not much variety. *Research Policy* 26, 141-156.

Penner-Hahn, J. and J. M. Shaver, 2005. Does international research and development increase patent output? An analysis of Japanese pharmaceutical firms. *Strategic Management Journal*, 26(2): 121-140.

Podolny, J., T. Stuart and M. Hannan. 1996. Networks, knowledge, and niches: competition in the worldwide semiconductor industry, 1984-1991. *American Journal of Sociology* 102: 659-689.

Porter, M. 1990. *The Competitive Advantage of Nations.* Free Press, New York.

Reagans, R. and E.W. Zuckerman, 2001. Networks, diversity, and productivity: the social structure of corporate R&D teams. *Organization Science* 12: 502-517.

Rodan, S. and C. Galunic, 2004. More than network structure: how knowledge heterogeneity influences managerial performance and innovativeness. *Strategic Management Journal* 25: 541-562.

Rosenkopf, L. & Nerkar, A., 2001. Beyond local search: Boundary-spanning, exploration, and impact in the optical disc industry. *Strategic Management Journal*, 22: 287-306.

Rosenkopf, L. and P. Almeida, 2003. Overcoming Local Search through Alliances and Mobility. *Management Science* 49(6). 0751-0766.

Saxenian, A.L. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128.* Harvard University Press, Cambridge.

Schumpeter, J., 1934. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Harvard University Press: Cambridge, MA.

Shaver, J.M. and F. Flyer. 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), pp. 756-770.

Singh, J., 2007. Asymmetry of Knowledge Spillovers between MNCs and Host Country Firms. *Journal of International Business Studies*, 38(5):764-786.

Song, J., P. Almeida and G. Wu, 2003. Learning by Hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management Science* 49(4): 351-365.

Sorenson, O. and J. A. C. Baum (Eds), 2003. *Geography and Strategy: Advances in Strategic Management,* Volume 20. Elsevier Science Ltd.

Sorenson, O., J.W. Rivkin and L. Fleming, 2006. Complexity, Networks and Knowledge Flow. *Research Policy*, 25: 994-1017.

Sorenson, O., J. Singh, 2007. Science, Social Networks and Spillovers. *Industry and Innovation,* 14(2): 219-238.

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

Subramaniam, M., N. Venkatraman, 2001. Determinants of transnational new product development capability: testing the influence of transferring and deploying tacit overseas knowledge, *Strategic Management Journal* 22(4): 359-378

Szulanski, G., 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal* 17: 27-43.

Teece, D.J., 1977. Technology transfer by multinational firms: the resource cost of transferring technological know-how. *Economic Journal* 87(346): 242-261.

Teece, D.J. 1986. Transaction cost economics and multinational enterprise. *Journal of Economic Behavior and Organization*, 7: 21-45.

Teece, G., Pisano, G. and R. Shuen, 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509-533.

Thompson, P. and M. Fox-Kean, 2004. "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment." *American Economic Review* 95(1): 450-460.

Thompson, P., 2006. Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations, *Review of Economics and Statistics* 88(2):383-389.

Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. *Rand Journal of Economics* 21(1): 172-187.

Trajtenberg, M., 2005. The Mobility of Inventors and the Productivity of Research. Mimeo.

Tushman, M. and L. Rosenkopf, 1992. Organizational determinants of technological change: towards a sociology of technological evolution. In *Research in Organizational Behavior*, Vol. 14, Staw BM, Cummings LL (eds). JAI Press: Greenwich, CT; 311-347.

Utterback, J.M., 1994. *Mastering the Dynamics of Innovation*. Harvard Business School Press: Boston, MA.

Williamson, O.E., 1985. *The Economic Institutions of Capitalism*. The Free Press.

Zander, I., 1997. Technological diversification in the multinational corporation: Historical evolution and future prospects. *Research Policy* 26: 209-227.

Zhao, M., 2006. Conducting R&D in Countries with Weak Intellectual Property Rights Protection. *Management Science*, 52: 1185 - 1199.

Table 1: Characteristics of the 1,127 firms included in the study

	Mean	Std Dev	Minimum	Maximum
Average geographic distance (in miles) between originating cities for patents arising in the last four years	638.4	639.6	0.0	4987.8
Herfindahl index of dispersion of recent patenting activity among locations	0.61	0.32	0.03	1.00
Number of locations where patents have originated in the last four years	12.1	15.2	1	132
Fraction of patents with first inventor having an address outside the US	0.44	0.46	0.00	1.00
Worldwide patent stock (4 yrs)	189.7	433.2	20	4581

Table 2: Variable definitions and summary statistics

		Obs	Mean	Std Dev	Min	Max
Dependent variable:						
value of innovation	The number of forward citations a patent gets.	529,669	4.82	7.43	0	779
Explanatory variables:						
R&D geographic spread	Natural logarithm of one plus the average value of geographic distance (in miles) between locations of the first inventors for all pairs of patents for the firm during the last four years.	529,669	6.24	1.07	0	8.56
R&D dispersion index	1 - Herfindahl of distribution of a firm's patents across different geographic locations. The index is between 0 and 1, with 0 being the lowest and 1 being the highest level of R&D dispersion	529,669	0.40	0.35	0	0.98
cross-regional knowledge sourcing	An indictor variable that is 1 if and only if at least one of the backward citations from the focal patent is to a patent originating in another location within the firm	516,497	0.16	0.37	0	1
cross-regional tie	An indictor variable that is 1 if and only if at least one of the inventors of the focal patent had a collaborative tie with someone from another region in the past four years	526,250	0.26	0.44	0	1
cross-regional move	An indictor variable that is 1 if and only if at least one of the inventors of the focal patent has moved from another region in the past four years	526,250	0.13	0.34	0	1
Control variables:						
team size	The number of inventors in the inventing team for the focal patent	528,318	2.47	1.68	1	34
technological breadth	1 - Herfindahl of NBER technological subcategories of the patent's backward citations. To avoid missing values, regression analysis sets this to 0 and defines a dummy *no backward citations* as 1 for patents with no backward citations	513,884	0.37	0.28	0	0.94
domain patent stock	The number of successful patents filed by the focal firm in the specific technological subcategory in the past four years	529,669	149.99	211.95	0	1,905
subsidiary patent stock	The number of successful patents filed by the focal firm from inventors in the focal region in the past four years	529,669	739.93	1,158.47	0	5,452
worldwide patent stock	The number of successful patents filed by the focal firm from inventors worldwide in the past four years	529,669	1,223.77	1,327.86	20	5,700
firm size	The total number of employees the firm had worldwide in the previous year (from COMPUSTAT). To avoid missing values, regression analysis sets ln(firm size) to 0 and defines a dummy *missing firm size* as 1 when value is missing.	199,100	121,663	161,243	21	813,400
R&D intensity	The ratio of the firm's R&D to sales in the previous year (from COMPUSTAT). To avoid missing values, regression analysis sets this to 0 and defines a dummy *missing R&D intensity* as 1 for firms with missing value.	205,285	0.07	0.28	0.00	28.31
region patent stock	Total number of all successful patents arising from inventors in the focal region in the past four years.	523,619	9,137.34	9,889.33	1	27,655
non-U.S. inventor	A dummy variable set to 1 if and only if at least one of the inventors of the focal patent resided outside the U.S.	529,669	0.55	0.50	0	1
IPR protection	Ginarte and Park (1997) index for intellectual property right protection in the country of the inventors - lowest value chosen in case of patents with inventors based in multiple countries.	525,334	4.22	0.43	0.33	4.86

Table 3: Correlation Matrix

		(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*	*(16)*
(1)	***value of innovation***	1.00															
(2)	***R&D geographic spread***	-0.01	1.00														
(3)	***R&D dispersion index***	0.01	0.78	1.00													
(4)	***cross-regional tie***	0.00	0.15	0.17	1.00												
(5)	***cross-regional move***	0.00	0.10	0.13	0.32	1.00											
(6)	***cross-regional knowledge sourcing***	0.04	0.19	0.23	0.16	0.13	1.00										
(7)	***ln(team size)***	0.04	-0.07	-0.14	0.32	0.19	0.03	1.00									
(8)	***technological breadth***	0.02	0.05	0.06	0.06	0.04	0.12	0.03	1.00								
(9)	***ln(domain patent stock)***	0.02	0.11	-0.02	0.06	0.02	0.14	0.13	-0.12	1.00							
(10)	***ln(subsidiary patent stock)***	0.00	-0.22	-0.41	-0.10	-0.12	-0.11	0.15	-0.05	0.47	1.00						
(11)	***ln(worldwide patent stock)***	-0.02	0.04	-0.01	0.00	-0.02	0.06	0.13	-0.06	0.69	0.61	1.00					
(12)	***ln(firm size)***	0.02	0.30	0.23	-0.02	-0.02	0.05	0.09	-0.03	0.45	0.35	0.76	1.00				
(13)	***R&D intensity***	0.01	-0.04	-0.05	0.02	0.01	-0.01	0.02	0.00	-0.01	-0.01	-0.05	-0.12	1.00			
(14)	***ln(region patent stock)***	-0.02	-0.29	-0.48	-0.12	-0.10	-0.17	0.15	-0.08	0.21	0.50	0.27	0.08	0.02	1.00		
(15)	***non-U.S. inventor***	-0.05	-0.27	-0.48	-0.17	-0.16	-0.15	0.14	-0.13	0.21	0.28	0.31	0.11	-0.01	0.55	1.00	
(16)	***IPR protection***	-0.05	0.17	0.30	0.10	0.12	0.10	-0.11	0.12	-0.11	-0.11	-0.18	-0.10	0.01	-0.35	-0.78	1.00

Note: These correlations are calculated using the subsample for which all of the above variables have values defined

Table 4: Negative binomial regression models (with firm fixed effects) analyzing the effect of distributed R&D on *value of innovation*

	(1)	(2)	(3)	(4)
R&D geographic spread	-0.049***	-0.045***		
	(0.002)	(0.002)		
R&D dispersion index			-0.146***	-0.250***
			(0.007)	(0.010)
ln(team size)		0.151***		0.148***
		(0.003)		(0.003)
technological breadth		0.086***		0.086***
		(0.005)		(0.005)
ln(domain patent stock)		0.024***		0.024***
		(0.001)		(0.001)
ln(subsidiary patent stock)		0.002		-0.000
		(0.001)		(0.001)
ln(worldwide patent stock)		-0.046***		-0.052***
		(0.003)		(0.003)
ln(firm size)		-0.001		0.008***
		(0.003)		(0.003)
R&D intensity		-0.046***		-0.050***
		(0.015)		(0.015)
ln(region patent stock)		0.015***		0.011***
		(0.001)		(0.002)
non-U.S. inventor		-0.207***		-0.207***
		(0.009)		(0.009)
IPR protection		-0.172***		-0.200***
		(0.008)		(0.008)
No backward citations		-0.019**		-0.019***
		(0.007)		(0.007)
Missing firm size		0.015		0.143***
		(0.038)		(0.038)
Missing R&D intensity		0.035**		-0.028*
		(0.016)		(0.016)
year dummies	included	included	included	included
technology dummies	included	included	included	included
Number of patent observations	529,655	522,697	529,655	522,697
Number of firms	1,113	1,109	1,113	1,109

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

Table 5: Negative binomial regression models (with firm fixed effects) analyzing the effect of cross-regional knowledge integration (and its interaction with *R&D geographic spread*) on *value of innovation*

	(1)	(2)	(3)	(4)	(5)	(6)
cross-regional knowledge sourcing	0.069***			0.048***		
	(0.004)			(0.005)		
cross-regional tie		0.024***			0.018***	
		(0.004)			(0.004)	
cross-regional move			0.009**			0.000
			(0.004)			(0.005)
R&D geographic spread	-0.046***	-0.046***	-0.045***	-0.048***	-0.048***	-0.047***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
R&D geographic spread X cross-regional knowledge sourcing				0.037***		
				(0.005)		
R&D geographic spread X cross-regional tie					0.016***	
					(0.004)	
R&D geographic spread X cross-regional move						0.020***
						(0.005)
ln(team size)	0.149***	0.145***	0.150***	0.149***	0.144***	0.150***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
technological breadth	0.080***	0.085***	0.086***	0.080***	0.085***	0.086***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ln(domain patent stock)	0.021***	0.023***	0.024***	0.021***	0.023***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(subsidiary patent stock)	0.004***	0.002**	0.002*	0.004***	0.002**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(worldwide patent stock)	-0.047***	-0.047***	-0.046***	-0.047***	-0.047***	-0.046***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln(firm size)	-0.001	-0.001	-0.001	-0.002	-0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R&D intensity	-0.047***	-0.046***	-0.046***	-0.048***	-0.046***	-0.046***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
ln(region patent stock)	0.016***	0.016***	0.015***	0.016***	0.015***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
No backward citations	-0.385***	-0.207***	-0.207***	-0.386***	-0.207***	-0.207***
	(0.022)	(0.009)	(0.009)	(0.022)	(0.009)	(0.009)
non-U.S. inventor	-0.165***	-0.164***	-0.170***	-0.166***	-0.164***	-0.170***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
IPR protection	-0.018**	-0.014*	-0.018**	-0.019***	-0.014**	-0.019**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Missing firm size	0.012	0.019	0.016	0.000	0.014	0.011
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Missing R&D intensity	0.042***	0.036**	0.035**	0.042***	0.036**	0.035**
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
year dummies	included	included	included	included	included	included
technology dummies	included	included	included	included	included	included
Number of patent observations	509,722	522,697	522,697	509,722	522,697	522,697
Number of firms	1,108	1,109	1,109	1,108	1,109	1,109

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

Table 6: Negative binomial regression models (with firm fixed effects) analyzing the effect of cross-regional knowledge integration (and its interaction with *R&D dispersion index*) on *value of innovation*

	(1)	(2)	(3)	(4)	(5)	(6)
cross-regional knowledge sourcing	0.074***			0.054***		
	(0.004)			(0.007)		
cross-regional tie		0.028***			0.020***	
		(0.004)			(0.004)	
cross-regional move			0.011**			0.005
			(0.004)			(0.006)
R&D dispersion index	-0.256***	-0.254***	-0.250***	-0.264***	-0.264***	-0.253***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
R&D dispersion index X cross-regional knowledge sourcing				0.073***		
				(0.018)		
R&D dispersion index X cross-regional tie					0.047***	
					(0.011)	
R&D dispersion index X cross-regional move						0.023
						(0.016)
ln(team size)	0.146***	0.141***	0.147***	0.146***	0.141***	0.147***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
technological breadth	0.080***	0.086***	0.086***	0.080***	0.086***	0.086***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ln(domain patent stock)	0.022***	0.024***	0.024***	0.022***	0.024***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(subsidiary patent stock)	0.002*	0.001	0.000	0.002	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln(worldwide patent stock)	-0.053***	-0.052***	-0.052***	-0.053***	-0.053***	-0.052***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln(firm size)	0.009***	0.009***	0.008***	0.008**	0.009***	0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R&D intensity	-0.051***	-0.050***	-0.050***	-0.051***	-0.050***	-0.050***
	(0.016)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)
ln(region patent stock)	0.012***	0.011***	0.011***	0.012***	0.011***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
No backward citations	-0.384***	-0.207***	-0.207***	-0.384***	-0.207***	-0.207***
	(0.022)	(0.009)	(0.009)	(0.022)	(0.009)	(0.009)
non-U.S. inventor	-0.193***	-0.191***	-0.197***	-0.195***	-0.192***	-0.198***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
IPR protection	-0.019**	-0.014*	-0.018**	-0.020***	-0.015**	-0.019**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Missing firm size	0.144***	0.150***	0.144***	0.133***	0.145***	0.142***
	(0.039)	(0.038)	(0.038)	(0.039)	(0.038)	(0.038)
Missing R&D intensity	-0.022	-0.028*	-0.028*	-0.019	-0.026	-0.028*
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
year dummies	included	included	included	included	included	included
technology dummies	included	included	included	included	included	included
Number of patent observations	509,722	522,697	522,697	509,722	522,697	522,697
Number of firms	1,108	1,109	1,109	1,108	1,109	1,109

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

Figure 1: Trends in Cross-Regional Knowledge Integration, 1986-1995