Multinational Firms and International Knowledge Diffusion: Evidence using Patent Citation Data

Jasjit Singh

Harvard Business School and Department of Economics

JOB MARKET PAPER

Nov 9, 2003

Abstract: This paper addresses three questions: (i) Are multinational firms (MNCs) really better than markets at transferring knowledge across borders? (ii) How actively do MNCs exchange knowledge with their host countries? (iii) Do they contribute as much to local knowledge as they learn from their host countries? To answer these questions, I analyze data on citations for over half a million patents from 4,400 firms and organizations from six countries, covering all manufacturing sectors. I estimate the probability of individual knowledge flows, as measured using patent citations, through a weighted maximum likelihood estimation approach for choice-based samples. Cross-border knowledge flows within the same MNC are found to be several times stronger than those between different entities even within the same country. Interestingly, these intra-MNC flows are equally strong in both directions between the home base and the foreign subsidiaries. Turning to intra-national knowledge flows, foreign MNC subsidiaries learn more from domestic entities than they contribute to host country knowledge, though this pattern differs across countries and industries. Knowledge flows from host countries to MNCs are in fact as strong as those between domestic entities, showing that MNC subsidiaries are not disadvantaged by their foreign affiliation. Finally, parent firms of MNCs with a higher fraction of innovative activity located abroad also learn more from other countries, suggesting that overseas innovation can increase an MNC's overall absorptive capacity for foreign knowledge.

Keywords: Knowledge Spillovers, Technology Diffusion, Multinationals, Foreign Direct Investment

JEL Codes: F2, L2, M2, O3, O5

E-mail: jsingh@hbs.edu. To be presented at *American Economics Association*'s 2004 Meetings. Comments welcome. Please check http://www.people.hbs.edu/jsingh for latest version. I am grateful to Bharat Anand, Richard Caves, Wilbur Chung, Ken Corts, Mihir Desai, Lee Fleming, Bronwyn Hall, Elhanan Helpman, Rebecca Henderson, Adam Jaffe, Tarun Khanna, Walter Kuemmerle, Josh Lerner, Megan MacGarvie, Anita McGahan, Marc Melitz, Jan Rivkin, Jordan Siegel, Olav Sorenson, Peter Thompson, Manuel Trajtenberg and seminar participants at Harvard, NBER, CCC Doctoral Consortium and AOM BPS Consortium for helpful comments and discussions. Errors remain my own.

1. Introduction

Who gains from globalization, and who loses? This is a topic of much debate, with the welfare effects of multinational firms (MNCs) being one of the most controversial issues. On the one hand, MNCs are dreaded as gigantic monopolies that wield more power than most countries and destroy local competition that comes in the way of their profits. At the other extreme, MNCs are viewed as benign carriers of resources and technology across borders, and as an efficiency-enhancing substitute for the market inefficiencies inherent in international economic transactions. Governments around the world continue to spend huge resources for attracting foreign direct investment (FDI), at least partly in the hope of positive externalities like knowledge spillovers.[1] The natural question that arises is: do MNCs really contribute as much to their host countries as they gain from them? I make this comparison in a specific setting – knowledge exchange between MNCs and their host countries. This is done through analysis of knowledge flows involving around 4,400 MNCs and domestic organizations in six major innovating countries – US, Japan, Germany, France, UK and Canada. The knowledge flows are measured using citations for over half a million patents from the period 1986-1995, covering all manufacturing sectors

Knowledge spillovers are often geographically localized.[2] The explanation offered is that knowledge has a "tacit" element, which cannot be easily codified for transmission over large distances (Arrow, 1969; Teece, 1977). However, an advantage of multi-nationality is thought to be the relative ease of knowledge transfer within firm boundaries even across countries (Kogut and Zander, 1993). This makes MNCs potentially important for international knowledge diffusion.[3] I first verify that MNCs are indeed better than markets at overcoming the difficulty of transmitting knowledge internationally. I then investigate the extent to which they exchange knowledge with domestic firms and organizations in their host countries. In

[1] Hanson (2001) summarizes evidence that most countries, both poor and rich, offer tax breaks and subsidies to MNCs as an incentive for investing. The attitude of policy makers is exemplified by the mission statement of Ireland's *Industrial Development Authority* (a government body for encouraging FDI): "We will win for Ireland, its people and its regions, the best in international innovation and investment so as to contribute to the continued transformation of Ireland to a world-leading society which is rich in creativity, learning and social well-being."
[2] Some studies that show this empirically are Jaffe, Trajtenberg and Henderson (1993), Audretsch and Feldman (1996), Keller (2002), Branstetter (2001), Jaffe and Trajtenberg (2002) and Peri (2002).
[3] FDI is not the only way in which global economic activity can contribute to international knowledge diffusion. Trade can also play an important role (Coe and Helpman, 1995; Eaton and Kortum, 1996; Keller, 1998; MacGarvie, 2003). However, the central interest of this paper is to study the role of foreign direct investment.

particular, I study how the extent of knowledge diffusion from MNCs to their host countries compares with the knowledge acquisition by MNCs through their overseas operations.

A large body of empirical work tries to find evidence of such knowledge spillovers by measuring productivity changes associated with FDI.[4] However, the results are inconclusive.[5] The challenge is in separating knowledge spillover effects of FDI from its competitive effects, and in dealing with heterogeneity in the quality of FDI. To avoid these confounding effects, some researchers have attempted to measure knowledge flows directly. Since citations between patents potentially help understand how an innovation builds upon existing ones, researchers often use them as a way of tracking knowledge flows. In this paper, patent citation data makes it possible for me to study *bilateral* micro-level knowledge flows between MNCs and their host countries. This approach captures knowledge flows leading to new innovations. This makes it less applicable for studying knowledge spillovers to developing countries, where the concern is often more to learn and apply existing technologies rather than to make major innovations at the world's technology frontier. For this reason, the current paper focuses only on six developed countries.[6]

While the international economics literature includes work on knowledge flows from MNCs to their host countries, it has paid little attention to the reverse phenomenon, i.e., knowledge acquisition as a possible reason for multi-nationality. However, anecdotal evidence suggests that MNCs often follow a deliberate strategy of locating their foreign subsidiaries in regions with high potential for learning from others.[7] Numerous studies by scholars of international business also emphasize the importance of such learning motives in

[4] Some papers in this spirit are Caves (1974), Lichtenberg and van Pottelsberghe de la Potterie (1998), Aitken and Harrison (1999), Baldwin, Braconier and Forslid (1999), Xu (2000), Girma and Wakelin (2001) and Haskel, Pereira, and Slaughter (2001). For recent surveys, see Blomström and Kokko (1998) and Moran (2001).

[5] Rodrik (1999) summarizes the situation as follows: "Today's policy literature is filled with extravagant claims about positive spillovers from direct foreign investment. These spillovers include technology transfer, marketing channels, superior management, and labor training… The hard evidence is sobering. Systematic plant-level studies from countries such as Morocco and Venezuela find little in the way of positive spillovers..." (p. 37).

[6] Even here, use of patent citations is not free from controversy. First, citations are meant more as a means to delineate the cope of a patent, and hence are a very crude measure of knowledge flows. Second, they capture only knowledge flows resulting in new innovations that further lead to patents. Therefore, research based on patent citations is best seen as a complement rather than substitute for alternative methods.

[7] In an HBS Case Study by Kuemmerle and Kobayashi, NEC (a Japanese firm) gave the following reason for opening an R&D facility in Princeton, USA: "The reason we chose Princeton was that it was located among regional clusters of scientific excellence, and it would be easy to absorb new sources of knowledge. There are many scientific institutes around Princeton - the Bell Labs, the IBM Research Institute, the SRI Institute, and so on."

overseas location choice by MNCs.[8] This has raised concerns that host countries might lose more from such "leakage" of domestic knowledge than they gain in the form of knowledge spillovers from FDI.[9] However, there are few empirical studies that directly measure the extent of micro-level knowledge flows to MNCs.[10] I fill this gap by measuring these flows and comparing them with knowledge flows from MNCs to domestic players.

Jaffe, Trajtenberg and Henderson (1993) pioneered the approach of using patent citation data to detect localized knowledge spillovers. They suggested comparing frequency of geographic co-location of citing and cited patents with that of control pairs with similar technological and temporal characteristics. Almeida (1996) was the first to apply this "matching" approach to study knowledge flows involving MNCs. However, he finds that some of the results (like evidence of knowledge spillovers from MNCs to local firms) are not robust, probably because of the small sample size of 228 patents. In addition, the specialized setting (semiconductors) makes the findings suggestive rather than conclusive. For example, my analysis involving all manufacturing sectors reveals that Almeida's finding of knowledge flows from domestic firms to MNCs *exceeding* those between domestic firms does not hold in the aggregate data. Frost (2001) applies the matching methodology in a more detailed study of knowledge flows, but his interest is only *unilateral* knowledge flows *to* foreign multinational subsidiaries. My paper instead focuses on calculating and comparing *bilateral* knowledge flows between MNC subsidiaries and local firms, as well as between MNC subsidiaries and their parent firm.

This paper also makes an important methodological contribution to the existing literature. In particular, a challenge with the matching approach is finding a good match along multiple dimensions, namely, the detailed technology class *and* the application date of patents. To overcome this challenge, I use a novel citation-level regression approach that estimates an equation for the probability of citation between two patents.[11] This approach gives the

---

[8] Examples are Hedlund (1986), Bartlett and Ghoshal (1989), Cantwell (1989), Porter (1990), Kogut and Chang (1991), Håkanson and Nobel (1993), Birkinshaw (1997), Kuemmerle (1999) and Chung and Alcacer (2002).
[9] Dalton and Shapiro (1995) summarize the debate in the US as follows: "Rapid growth of foreign R&D in the US has led to concerns about an erosion of US technology leadership, and about the clustering of foreign R&D centers around major US research universities that receive funding from federal grants and other taxpayer support. Some observers have questioned the quality of the research effort by foreign companies... They have argued that US research centers of foreign companies are merely 'listening posts' that focus on technology scanning."
[10] As discussed later, some exceptions are Almeida (1996), Frost (2001) and Branstetter (2002).
[11] For comparability with existing matching-based studies, and as a robustness check, I also present my results using a matching approach. Unlike existing matching studies, however, my matching-based analysis uses a more detailed

flexibility of introducing detailed controls for the technological, temporal and geographic characteristics of patents. I use choice-based sampling (i.e. over-sampling on the actual citations or the "ones") to be able to estimate the model despite the small fraction of "ones" in the population. A weighted maximum likelihood (WESML) approach is used to circumvent biases from such a sampling approach (Manski and Lerman, 1977).[12] I control for technological relatedness of patents at the detailed 9-digit technology level, while taking into account not just the primary technological classification of patents but all technological classes listed for a patent. As Thompson and Fox-Kean (2003) show, absence of such detailed controls could have led to large biases in previous studies.

My analysis reveals strong cross-border within-MNC knowledge flows, which are several times as large as those between different entities even within the same country.[13] The extent of these intra-MNC knowledge flows is found to be of comparable magnitude in either direction between the home base and the overseas subsidiaries. This goes against the conventional economic view that MNCs transmit knowledge only from the headquarters to the foreign divisions. Turning to intra-national knowledge flows, significant bi-directional knowledge diffusion is found between domestic organizations and MNCs. There is little evidence that MNC subsidiaries face a "liability of foreignness" (Hymer, 1976), i.e., are not able to tap into the localized knowledge exchange in a country. In fact, the intensity of knowledge diffusion from domestic organizations to MNC subsidiaries is comparable to that among domestic organizations. On the other hand, diffusion from multinationals to domestic organizations turns out to be somewhat weaker in the aggregate analysis. These results are found to vary across industries and countries in a manner consistent with existing evidence on overseas subsidiary location by firms that are technological laggards or leaders in their fields (Chung and Alcacer, 2002). I also find that multinational subsidiaries are particularly good at

---

9-digit technological classification of patents in order to avoid biases that would result if only a coarse 3-digit classification were used (Thompson and Fox-Kean, 2003).

[12] Papers that use matched samples also sometimes include regression analysis based on these samples to try to correct for imperfect matching. However, these regressions do not take into account the implicit differences in sampling rates between actual and unrealized patent citations.

[13] In a related paper (Singh, 2003), I explore inventor mobility and social networks as mechanisms for knowledge diffusion, and as one possible explanation for geographic localization of knowledge spillovers. In analysis not included in that a paper, I found that higher intra-MNC knowledge flows are also at least partly a result of higher cross-border mobility of inventors within an MNC, and of stronger direct or indirect collaborative ties between inventors in different international divisions of the same MNC.

learning from each other, a finding consistent with Porter (1990) and Feinberg and Gupta (2003).

To further explore the link between multi-nationality and international knowledge diffusion, I also study the direct effect of FDI on cross-border patent citations between different entities. Globerman, Kokko and Sjöholm (2000), using data on 220 Swedish patents, find higher outward FDI but not inward FDI to be associated with higher patent citations to the partner countries. Branstetter (2000) studies effects of Japanese FDI in the US using a firm-level patent citation dataset, and finds learning effects from both inward FDI and outward FDI. While he measures FDI using simply the number of subsidiaries, I extend his work by using the number of patents arising from a firm's foreign subsidiary as a measure of its absorptive capacity for knowledge in that country (Cohen and Levinthal, 1989). I find that, as the innovative output of an MNC's foreign subsidiary increases, the bilateral patent citation rate between the MNC's home base and the other country also increases. This provides additional evidence that innovation-centered FDI is a channel for international knowledge diffusion.

The rest of the paper is organized as follows. Section 2 describes the patent citation data and the parent-subsidiary data for MNCs. Section 3 takes a matching-based approach to measure knowledge flows among MNCs and domestic organizations. Section 4 studies the same issue using a citation-level regression framework based on choice-based sampling, and explores cross-country and cross-technology differences in the aggregate pattern of knowledge flows. Section 5 investigates whether innovative activity in an overseas subsidiary also increases the ability of the home base of an MNC to acquire knowledge from other firms and organizations in the foreign country. Section 6 justifies the use of data from US Patent Office (USPTO) in measuring international knowledge diffusion by offering a robustness test using a sample of European Patent Office (EPO) patents. Section 7 offers concluding thoughts.

2. Data on Patent Citations and Multinational Ownership

I use citations made by patents as evidence of knowledge flows across innovations. Since patent citations are meant to delimit the property rights of the citing patent by listing all relevant "prior art", they leave behind a trail of how new innovations build upon existing ones. As opposed to citations in academic papers, the inventor has an incentive not to include unnecessary citations, since doing so only reduces the scope of his or her patent. The inventor is also legally bound to state all relevant "prior art" for the innovation, with the possibility of legal prosecution if found to deliberately omit citation of prior work on which he or she builds. An important job of the patent examiner is to ensure that all relevant prior art is cited, increasing the objectivity of patent citations. This means, however, that a patent examiner might add citations to patents of which the original inventor was not even aware. This adds noise to patent citations as a measure of knowledge flows. As recent studies based on comparison of citation data with inventor surveys have shown, the correlation between patent citations and actual knowledge flows is still fairly high (Jaffe, Trajtenberg and Fogarty, 2000; Duguet and MacGarvie, 2002), justifying the use of patent citations as a reasonable proxy for knowledge flows in large-sample studies.[14]

2.1. Patent Citation Data

Patents from different patent offices are not comparable to each other because of different patent breadth, patenting costs, approval requirements, citation practices and enforcement rules in different countries. Therefore, it is common practice to use data from a single patent granting country like the US (Jaffe and Trajtenberg, 2002), the UK (Lerner, 2002) or Sweden (Globerman, Kokko and Sjöholm, 2000) in order to standardize the measure of innovation for research purposes. Following this practice, I use a data set on US patents, constructed by merging data from the US Patent Office with an enhanced version from NBER (Hall, Jaffe and Trajtenberg, 2002). Since the US is the largest and technologically most advanced market in the world, an innovation made with a global market in mind is likely to be patented in the US. This makes patent data from US Patent Office a natural choice to study.

I used the address of the first inventor to determine the country where the innovation took place.[15] In order to ascertain whether a patent originated from a domestic organization or from

[14] Firms rely not just on patents but also on other mechanisms such as secrecy, complementary sales, service capabilities, and quicker lead times for appropriating returns from innovation (Levin, Klevorick, Nelson and Winter, 1987). In addition, when patents do get used, it is sometimes to block the development of a substitute or a threat to force rivals into negotiations. Nevertheless, patents have been shown to be a reasonable, though noisy, measure for innovation and are widely used as such in research (Griliches, 1990). It is worth emphasizing that I *do not* use raw patent or citation counts for cross-country comparisons, so there is no bias from differences in *propensity to patent* across different types of inventors. I only use US patent data to track knowledge flows *conditional on* the given set of innovations as embodied in US patents.

[15] The share of potentially controversial cases, where patents involve inventors from multiple countries, is less than 2% in the data.

the local subsidiary of a foreign MNC, I checked whether the "home country" of the assignee organization was the same as the country of innovation.[16] This turned out to be a non-trivial task, and is explained in more detail in the following subsection. I restricted my analysis to patents with application dates between 1986 and 1995 since my determination of the assignees' corporate parents is based on a definition from around 1990 (as explained below), making mergers and acquisitions a serious concern when using very long time windows.[17]

This paper restricts itself to six leading economies, since the number of patent citations arising from the remaining countries is smaller and hence statistically less useful. The included countries are: U.S., Japan, Germany, France, U.K. and Canada. In Table 1, column (1) shows that the number of patents from these countries for the period 1986-1995 is about 0.9 million, or about 91% of all USPTO patents. As column (2) indicates, about 83% of the patents are owned by firms or organizations rather than individuals. I drop all unassigned patents since our interest here is on inter-organizational knowledge flows.

One might be concerned about possible biases from using US patent citation data for cross-country comparisons. To avoid these biases, we need to control for differences across inventors in their propensity to cite US patents. For example, inventors based in European countries might cite the European Patent Office (EPO) patent corresponding to an innovation instead of citing the USPTO patent. The regression models in this paper include country fixed effects to take care of such issues. Likewise, fixed effects for industries and time periods account for systematic differences in propensity to cite along these dimensions.

2.2. Multinational Data

A crucial step in building the dataset was identifying whether an assignee firm, as listed in the patent data, had its home base in the country of patenting or if it was an affiliate of a foreign MNC.[18] The patent database lists about 175,000 assignees, and it is difficult to match these assignees to their parents. For example, there is no systematic rule as to whether patents

[16] Note that this means that the innovations made by an MNC in its home country are a part of "innovations arising from domestic organizations" in that particular country, and the innovations made by an MNC in countries other than its home country are a part of "innovations arising from subsidiaries of foreign MNCs".
[17] An earlier version of this paper reported results for a longer time frame (1980-1999) and four additional countries with fewer US patents – Switzerland, Italy, Netherlands and Sweden. The qualitative results there were similar.
[18] Some of the "multinationals" are actually not firms but government organizations like NASA and the US Armed Forces. However, the relative role of such organizations in US patents arising overseas is quite small, so I do not make that distinction in the analysis.

originating from IBM researchers based in a German subsidiary are listed using the same assignee code as the US-based parent "IBM" or a separate assignee "IBM Germany" (or another name from which it is even harder to infer if the assignee is a subsidiary of IBM).[19]

Appendix A gives the details of the extensive data cleaning exercise for a subset of the assignees, which led to identifying about 4,400 unique organizations, including about 3,000 firms, 400 government-affiliated bodies, 550 research institutes and 450 universities worldwide. Column (3) of Table 1 shows that these assignees account for about 556,000 of the patents coming from the six countries considered. This comprises about 73% of all patents owned by organizations for this set of countries. The remaining patents are spread among about 165,000 assignees not inspected above. Since I cannot be sure of the home country for these assignees, I drop data for these assignees from the analysis reported in this paper.[20] Column (4) of Table 1 shows the extent of patenting activity arising from foreign MNC subsidiaries in each country. This fraction is about 9% of the entire data, though it varies across countries from as low as around 2% for Japan to as high as about 50% for Canada. Although the variation in the extent of foreign MNC patenting activity is interesting in itself, this paper focuses only on the extent of knowledge diffusion (as measured by patent citations), taking the existing distribution of patenting activity as given.

3. A Matching Approach for Measuring Knowledge Diffusion

In using patent citations to study knowledge diffusion, we need to ensure that the observed citation pattern is not merely a reflection of the geographic distribution of patenting activity. For example, since innovations in the same or related industries are likely to be located in the same region (often for reasons other than just knowledge spillovers), a higher citation frequency between patents in the same region need not be evidence of stronger localized knowledge diffusion. Jaffe, Trajtenberg and Henderson (1993) suggest an approach that tries to

[19] I defined the subsidiary as being an entity in which an MNC has a majority stake. While even a lower stake can give a multinational enough control over a foreign company, I wanted to avoid the situation in which a company could not be identified with a unique parent. For cases where two multinationals had exactly 50-50 stake in a company, I broke the tie by assuming it was a part of the MNC whose name appears first in the joint venture name.
[20] For robustness, I also tried to keep these observations and computed the "home country" for the remaining assignees as the country in which the largest fraction of its patents originate. When a company patent uses a single assignee number, this approach would correctly identify the home country in most cases since MNCs do most of their research at home. In cases where different divisions of the same firm use different assignee numbers and have

take this into account by defining the appropriate benchmark as being the citation probability from the original patent to a random patent with similar technological characteristics as the originally cited patent.

3.1. Details of Matching

In order to find a control patent for any given patent, I looked for another patent with exactly the same *9-digit* primary classification as defined by the US Patent Classification (USPC) system.[21] Since the time lag between two patents is also an important determinant of the probability of citation, my matched sample only includes those cited patents for which a control patent could be found with the application year being within one year of the original. Since our interest is in inter-organizational flows, all citations for which either the original or the control involved a self-cite from a firm or organization to itself were excluded from further analysis. These steps led to dropping about half of the citations from the original data.

Figure 1 shows how we can test for knowledge diffusion involving MNCs. As the top part of the picture shows, if the fraction of MNC patents (i.e. patents originating from local subsidiaries of foreign MNCs) is higher in the set of patents actually cited by domestic organizations than in the set of control patents corresponding to the cited patents, it is evidence in favor of knowledge diffusion from MNC subsidiaries to domestic organizations (which I call multinational-to-domestic, or simply M→D diffusion)[22]. The t-statistic used to test if the two fractions are statistically different is calculated as

$$t_{M\to D} = \frac{p_{M\to D} - p'_{M\to D}}{\sqrt{\dfrac{p_{M\to D}(1-p_{M\to D})}{N_D} + \dfrac{p'_{M\to D}(1-p'_{M\to D})}{N_D}}}$$

where $p_{M\to D}$ is the ratio of number of citations from domestic organizations to MNC entities ($n_{M\to D}$) to the total number of citations N_D (= $n_{D\to D} + n_{D\to M} + n_{D\to F}$) made by domestic

---

R&D bases in different countries, there will errors in classification. Nevertheless, following this empirical strategy led to results qualitatively similar to those reported in the paper.

[21] Most existing studies match only using a *3-digit* or an even coarser classification for the matching. Since there are only 410 categories of technology in the 3-digit classification, a 3-digit matching leaves room for tremendous heterogeneity even within a technology class, biasing the results. In contrast, there are around 150,000 subclasses of technology in the 9-digit classification, leading to control patents with a much better match with the original patents.

[22] Note that if a patent in set A cites another patent in set B, then the direction of knowledge flow is from B to A, i.e., opposite from the direction of citation. As a convention, I use the arrow to indicate the direction of knowledge flow, not the direction of citation.

entities. I similarly compute the t-statistics to test for domestic-to-domestic (D→D) diffusion, domestic-to-multinational (D→M) diffusion and multinational-to-multinational (M→M) diffusion.

3.2. Results from Matching

Table 2 reports the analysis of localized knowledge diffusion among domestic organizations (D→D flows). Column (1) gives the total number of citations made by domestic organizations, and columns (2) and (3) respectively give the number and fraction of these that were made to patents by domestic organizations. Columns (4) and (5) report the same analysis for patent dyads obtained by replacing each original cited patent by its control patent. Column (6) reports the difference of the proportions from columns (3) and (5), column (7) calculates the t-statistic for testing the equality of the two proportions, and column (8) gives the ratio of the two proportions (which I call the D→D diffusion index).[23] The overall D→D diffusion index of 1.22 indicates that patents from two domestic firms are 22% more likely to cite each other than are two random patents with similar technological and temporal characteristics.

Table 3 turns to analysis of localized knowledge diffusion from domestic organizations to local subsidiaries of foreign MNCs (D→M flows). This time, column (1) gives the total number of citations made by subsidiaries of foreign MNCs, and columns (2) and (3) respectively give the number and fraction of these made to patents by domestic organizations, with columns (4) and (5) doing the same analysis for the matched control patents. The localization result discussed above for D→D flows is shown to exist even for D→M flows, with the probability of such citations exceeding that of control patents by an almost identical figure of 20%.

In Table 4, a similar methodology is used to calculate knowledge diffusion from local subsidiaries of foreign MNCs to domestic organizations (M→D flows). While the localization of knowledge diffusion result still holds, the extent of knowledge diffusion and the t-values are smaller than in the earlier cases. The magnitude of the M→D index (1.13) is smaller than

[23] I could also have used the *difference* between the two proportions itself as a measure of the strength of spillovers. That, however, introduces scale effects that make it impossible to make comparisons across countries or across different types of spillovers. For example, if the fraction from actual data is 0.10 and that from control data is 0.05, the strength of the spillovers should probably be interpreted as being larger than it would be if the two fractions were

in the D→D case (1.22) or D→M case (1.20). This suggests an asymmetry in knowledge flows between multinational firms and domestic organizations: Even though there are knowledge flows in both directions between the two, MNC subsidiaries learn more from the domestic organizations than they contribute towards learning by domestic organizations. I postpone a formal statistical test of this claim until the next section.

Finally, Table 5 calculates knowledge diffusion between local subsidiaries of different foreign MNCs (M→M flows), and finds the magnitude (1.23) to be at least as strong as D→D case (1.22) or D→M case (1.20). Thus multinational subsidiaries learn not just from the other domestic organizations, but also from each other.

4. Citation-Level Regressions

As mentioned above, control patents that match well along multiple dimensions are very hard to find. If we drop the observations where no reasonable matches are found, the results might be biased if the dropped observations are systematically different from the rest.[24] Even for citations for which we have reasonable matches, matching is rarely perfect when several dimensions are important. For example, most patents have not just "primary" but also "secondary" technology classification, with what is primary versus secondary not necessarily being a true reflection of a patent's characteristics. Technological relatedness of patents might therefore show up as overlap along any of their subclasses, and not necessarily as equality of the primary one. This is very difficult to take into account in the matching approach.[25]

0.50 and 0.45 respectively. A *ratio* of the proportions captures this, but a *difference* of the proportions fails to capture this.

[24] Thompson and Fox-Kean (2003) illustrate that the bias could go either way. For example, technologically active industries produce a lot of patents, and hence have a higher probability of finding control patents within a narrow time window. But if they are also more likely to have localized spillovers, the overall results will be biased upward. On the other hand, consider the subclasses that are narrowly defined. Since that also have fewer potential control patents, observations from these subclasses have a lower probability of being included. But these subclasses are also likely to have less unobserved heterogeneity, and therefore are more likely to be geographically concentrated. The higher omissions from these subclasses would lead to a downward bias in overall results on localized knowledge spillovers.

[25] Thompson and Fox-Kean (2003) give the following example to show how ignoring the secondary classification could lead to a situation where the originating and citing patents come from the same industry but the control and citing patents are essentially unrelated: There is a citation between patent number 6,186,059 (with primary class 100 and a secondary class 426) and patent number 5,064,667 (with primary class 426) arguably because these two patents have an overlapping class 426 (which pertains to "Food or Edible Material"). A control patent based only on the primary class will miss this, and hence be less likely to make a similar citation.

To overcome these issues, I use a citation-level regression framework to model the probability of citation between two patents. Imagine that the probability that a patent *K* cites a patent *k* is given by a "citation function" P(*K*, *k*). My interest lies in studying how P(*K*, *k*) differs when one or both of these patents arise from MNC subsidiaries rather than domestic firms within the same country, or from the same MNC or different entities in two different countries. Among the explanatory variables, I include dummy variables for all dimensions along which we would have ideally liked to do the matching. This gives the flexibility of using multiple control variables to better control for propensity to cite even in cases where good matches do not exist.

4.1. Choice-Based Sampling [26]

Since the number of potentially citing and cited patents can be of the order of a million, the number of all possible dyads (K, k) can be of the order of a trillion. In principle, we could take a random sample of patent dyads from the population of all possible dyads. We could then define a binary variable *y* that equals 1 if the citation actually takes place, and 0 otherwise, and estimate the citation function by assuming that it can be approximated using a logistic functional form. In other words, the dichotomous dependent variable *y* would be taken as a Bernoulli outcome that takes a value 1 for observation *i* with the probability

$$\Pr(y = 1 \mid x = x_i) = \Lambda(x_i \beta) = \frac{1}{1 + e^{-x_i \beta}}$$

where $\mathbf{x}_i$ is the vector of covariates and β is the vector of parameters to be estimated. However, an estimation approach based on random sampling of patent pairs is not practical because actual citations are very rare. There are only about seven actual citations for every one million potential citations, making estimation impossible even with large samples. From an informational point of view, it would be desirable to have a higher fraction of observations with $y = 1$ in the sample.[27]

[26] Because of similar methodology, the discussion here overlaps significantly with that in Singh (2003).

[27] The asymptotic covariance matrix for the MLE for logit is given by $\left[\sum_{i=1}^{n} \Lambda_i (1 - \Lambda_i) x_i x_i^{'}\right]^{-1}$ (see Greene, 2003, p. 672). If the logit model has some explanatory power, Λ_i is larger (i.e. closer to 0.5 for rare events) when $y_i = 1$. Thus $\Lambda_i(1-\Lambda_i)$ is larger for these observations, implying that having a higher fraction of 1's in the sample would, other things being equal, reduce variance.

An alternative approach is to use a "choice-based" sampling procedure that deliberately oversamples the patent pairs with actual citations, i.e., with $y = 1$.[28] In this approach, the sample is formed by taking a fraction α of the population's dyads with $y = 0$, and a fraction γ of the dyads with $y = 1$, where α is much smaller than γ. Since this stratification is done on the dependent variable, however, using the usual logistic estimates would lead to a selection bias. If we were sure that the underlying functional form is indeed logistic, it is possible to make a correction for this bias.[29] The efficiency of the correction, however, depends crucially on the assumption that the logit functional form is not misspecified (Manski and Lerman, 1977; Cosslet, 1981). Since there is no reason to expect that the functional form would hold exactly, a better option is to use the *weighted exogenous sampling maximum likelihood* (WESML) estimator suggested by Manski and Lerman (1977). The central idea is to explicitly recognize the difference in sampling of 0's and 1's by weighting each term in the log likelihood function by the inverse of the ex ante probability of inclusion of the corresponding observation in the sample. In other words, each sample observation is weighted by the number of elements it represents from the overall population in order to make the choice-based sample "simulate" a random exogenous sample.[30] The

---

[28] For a general discussion of choice-based sampling, see Amemiya (1985, pp. 319-338), Greene (2003, p. 673) or King and Zeng (2001). Sorenson and Fleming (2001) have also used this technique for predicting patent citations.

[29] The probability of a citation *conditional on the dyad being in the sample* flows from Bayes' rule:

$$\Lambda_i^{'} = \frac{\gamma \Lambda_i}{\gamma \Lambda_i + \alpha(1-\Lambda_i)} = \frac{\gamma}{\gamma + \alpha e^{-\beta X_i}} = \frac{1}{1+e^{-(\ln\left(\frac{\gamma}{\alpha}\right)+\beta X_i)}}$$

This differs from Λ_i as there is now an extra term $ln(\gamma/\alpha)$ in the exponent, leading to a bias. However, since the functional form is still logistic, a simple estimation strategy is to use the usual logit-based maximum likelihood estimation, and then subtract $ln(\gamma/\alpha)$ from the estimate for the constant term.

[30] For intuition, let the joint probability distribution of x and y be g(x,y) be for the sample, and g*(x,y) for the population. Let the fraction of elements with $y = j$ be f(j) in the sample, and f*(j) in the population. Let n and N be sample size and population size respectively, and n_j and N_j be the number of these with $y = j$. Using conditional probability rules,

$$g(x,j) = \Pr(x \mid y=j) f(j) = \frac{g^*(x,j) f(j)}{f^*(j)} = \frac{g^*(x,j)(n_j/n)}{N_j/N} = \frac{N/n}{w(j)} g^*(x,j)$$

where $w(j) = N_j/n_j$ is the reciprocal of the sampling rate for observations with $y = j$. Let P(y_i) be the probability of $y = y_i$ conditional on $x = x_i$ in the population. Then, the expected value of the weighted likelihood function is

$$E \ln L_w = \int \left(\sum_{i=1}^{n} w(y_i)[\ln P(y_i)] \right) g(x,y_i)dx = \sum_{i=1}^{n} \left(\int w(y_i)[\ln P(y_i)] \frac{N/n}{w(y_i)} g^*(x,y_i)dx \right) = \frac{N}{n} \int \left(\sum_{i=1}^{n} [\ln P(y_i)] \right) g^*(x,y_i)dx$$

Thus, ignoring the constant scaling factor N/n, the expected value of the weighted log likelihood equals the expected log likelihood for the same sample resulting through random exogenous sampling from the population. As shown formally in Amemiya (1985, section 9.5.2), this ensures consistency of WESML estimation.

WESML estimator is then obtained by maximizing the following weighted "pseudo-likelihood" function:

$$\ln L_w = \frac{1}{\gamma}\sum_{\{y_i=1\}}\ln(\Lambda_i) + \frac{1}{\alpha}\sum_{\{y_i=0\}}\ln(1-\Lambda_i) = -\sum_{i=1}^{n} w_i \ln(1+e^{(1-2y_i)x_i\beta})$$

where $w_i = (1/\gamma)\, y_i + (1/\alpha)(1-y_i)$. In addition, the appropriate estimator of the asymptotic covariance matrix is White's robust "sandwich" estimator used for pseudo-maximum likelihood estimation. Since the same citing patent can occur in multiple observations, the standard errors should be calculated without assuming independence across these observations.

4.2. Sample Construction

While the above choice-based sampling design is an improvement over random sampling, it can be improved further. In particular, the above approach still samples all *y = 0* observations with equal probability, irrespective of their "relevance." To see why that might be an issue, recall that the technological similarity of two patents is particularly relevant for the probability of citation. Therefore, to estimate other coefficients efficiently, we would ideally like to have a sample with sufficient variation in the dependent variable within groups of patents with comparable degree of technological similarity. However, randomly drawn *y = 0* observations typically have patent dyads that are not technologically related, while *y = 1* observations are very often technologically related. One way to improve the estimation is to extend the basic choice-based sampling design to also allow matching on the technological class of the citing and cited patents.[31] The weighted likelihood function described above now has to be generalized since the probability of a *y = 0* pair getting selected now depends on the technological classes of the two patents. As Appendix B shows, this generalization involves defining the weight attached to a *y = 0* observation as the reciprocal of the ex ante probability of a *y = 0* population pair *with the same respective technological cell* (i.e. combination of technological classes for the citing and cited patents) being selected into the sample.

I define the population of possible citations as all pairs of citing and cited patents in my data (over half a million patents from 1986-1995) such that the citing year does not come before the cited year. The sample used in regression analysis was drawn from this population as follows: First, all actual citations (i.e., elements with $y = 1$) were included in the sample, except for self-citations from a geographical division of an organization to itself. Each of these "ones" was then matched with multiple "zero" citations (i.e., patent pairs with $y = 0$) that have the same "cell" as defined by the characteristics of the actual citation.[32] This was done while making sure that no self-citation from a geographical division of an organization was included among the control citations either. At least one control citation was chosen from even those cells that did not have an actual citation. This led a sample of a little over half a million actual and potential citations.

4.3. Control Variables for Propensity to Cite

In testing the impact of the variables of interest on the probability of citation between two patents, we need to control for other factors that might affect the propensity to cite between the two. The first issue is time effects. The citation probability is known to increase initially as the time lag between the citing and cited patents increases, but then starts to fall after a certain point (Jaffe and Trajtenberg, 2002). Since I am not interested in estimating the exact shape of this lag function but in only controlling for the effect, I introduce dummy variables for the number of years of lag between the citing and cited patent. In addition, since the patent citation rate may change over time, additional dummy variables are used to capture the citing year fixed effects.

The next issue is that innovators in different countries might have a different propensity to cite patents registered with the US patent office. For example, a US patent filed by a Japanese firm might not necessarily cite the US patent for the innovation it builds upon, but might instead cite the corresponding Japanese patent for that innovation. In order to avoid possible biases arising from this, all regressions include citing country fixed effects as mentioned earlier. Section 6 carries out additional robustness checks to ensure that MNCs and domestic firms *within the same country* also do not differ in their propensity to cite USPTO

[31] While the "matched samples" are often constructed to be of the same size as the original set, it need not be so. In fact, given that WESML robust standard errors tend to be quite large when α is small, it is often advisable to use larger control sets to improve precision, if computing resources permit (King and Zeng, 2001).
[32] Generalizing on the above discussion, this matching was actually done not just along the technological characteristics but also included the citing country to improve individual country estimates for countries with lesser

patents in a way that could lead to incorrect inference. In addition, since patents in different industry categories have different propensities to cite others, I include fixed effects for the broad technological category of the citing patent, as defined in the NBER database.

An important contribution of this paper is the use of better controls for technological relatedness of the citing and cited patents. Patents with similar technological characteristics are more likely to cite each other. In order to capture technological similarity between two patents, I define dummy variables for the same broad technological category, the same technological subcategory, the same 3-digit primary class and the same 9-digit primary class. This gives a set of hierarchical controls for technological relatedness of patents, something typically missing in earlier papers using patent citations to measure knowledge diffusion. Further, the designation of a subclass as being primary for a patent can often be almost arbitrary. Therefore, I also include another dummy variable that captures whether at least one of the secondary subclasses of a patent is the same as the primary or one of the secondary subclasses for the other patent in the pair. As Thompson and Fox-Kean (2003) have shown, absence of such detailed controls for technological relatedness has caused large biases in previous studies measuring knowledge spillovers. [33]

4.4. Results

Table 6 gives a summary of all variables used in the regressions below. The dependent variable for all regressions will be a binary variable that is 1 for patent pairs with a citation, and 0 for pairs with no citation. All regressions include fixed effects for technological category of citing patent, country of citing patent, citing patent year and time lag between

---

patenting. I chose a higher number of control citations per actual citation for countries with smaller number of citations. The weights of the "zeroes" from different countries were adjusted accordingly.

[33] Existing regression-based studies do not typically capture knowledge flows at the individual citation-level. Instead, they use a more aggregate dependent variable like the number of citations at the level of a country (e.g., Globerman, Kokko and Sjöholm, 2000), a firm (e.g., Branstetter, 2000) or some other level of aggregation (Jaffe and Trajtenberg, 2002). These models control for propensity to cite by including a measure of "average technological distance" between the aggregate sets of citing and cited patents, and ignore within-set heterogeneity. Since this technological distance is measured only at the coarse 2 or 3-digit technology classification level, the issue of biases discussed earlier still remains. To see this, imagine that there are no localized knowledge spillovers, and that technological distance as measured at the 3-digit level is the same for all observations. Even so, citing and cited sets of patents with a higher fraction of patent pairs belonging to the same detailed 9-digit technology will have a higher fraction of citations and also a higher co-location of citing and cited patents (the latter resulting purely from technological specialization of regions), leading to a spurious correlation.

patents.[34] All tables will report robust standard errors to account for the weighting in the likelihood function. In addition, these standard errors are calculated without assuming independence of observations from the same citing patent.

The initial analysis, as reported in Table 7, addresses the following questions: (1) How much stronger are cross-border knowledge flows within an MNC than outside? (2) How strong is the evidence on geographic concentration of knowledge spillovers, when measured at a country level? Column (1) reports the results without controlling for technological relatedness of the citing and cited patents. The positive and significant coefficient on the dummy variable *within same MNC*, which has an estimated positive and statistically significant value of 3.29, suggests that cross-border knowledge flows are stronger within the same MNC than between two different organizations. To understand the economic magnitude of intra-MNC knowledge flows, we need to look at the reported marginal effect that appears in square brackets (after multiplying by a million for readability).[35] This tells that there is 18.8 in a million higher citation probability for knowledge flows within the same MNC than between different organizations in different countries. Given that the predicted citation rate between two random patents is about of 5.7 in a million, this suggests that patents from different divisions of an MNC are more than four times as likely to have a citation than are patents from different organizations from different countries. Similarly, the positive and significant coefficient of 0.67 for the variable *within same country* suggests that being in the same country increases inter-organizational patent citation rate by about 3.8 in a million, or about 67%, over the reference category.

The above results may, however, at least partly reflect the fact that patents within the same country or within the same MNC may have a higher propensity to cite each other simply because they are more related in their technology. This concern is proven to be valid in column (2), which shows that including controls up to the 3-digit primary technological class of the patents reduces the effects of *within same country* and *within same MNC* somewhat.

[34] The estimates for these fixed effects were found to be broadly consistent with the results reported in Jaffe and Trajtenberg (2002) using non-linear least squares estimation of a citation function. My goal here, however, is only to control for these effects while exploring other questions of interest for this paper, so I do not discuss these effects in detail.

[35] Recall that $\Lambda(\boldsymbol{x\beta})$ gives the probability of citation. The marginal effect of a variable j is therefore given by β_j $\Lambda'(\boldsymbol{x\beta})$. From the logit form, it is easy to show that this equals $\beta_j\ \Lambda(\boldsymbol{x\beta})[1-\Lambda(\boldsymbol{x\beta})]$. One can then substitute either the mean predicted probability or simply the population mean for $\Lambda(\boldsymbol{x\beta})$ for getting an estimate of the marginal effect. I report the former. The latter estimate is typically of the same order of magnitude but slightly higher in exact value.

This corresponds to the technological controls used in most existing studies. However, Thompson and Fox (2003), in their criticism of the existing literature, have shown that controls at just the 3-digit technological class level are not sufficient, and can lead to biases.

Columns (3) and (4) include more detailed technological controls to address the above criticism. As column (3) shows, controlling for the detailed 9-digit technological classification does reduce the estimates of the two effects. Similarly, column (4) shows that controlling for overlap of secondary technological classes reduces the two estimates even further. However, the two estimates are still statistically and economically quite significant. In particular, the estimates suggest that patents from different subsidiaries of the same MNC are around three times as likely to have a citation than are patents from different organizations in different countries. Likewise, patents from different organizations within the same country are about 58% more likely to have a citation than are patents from different organizations in different countries, even after stringent controls for technological relatedness. Interestingly, the probability of citation between patents from different international divisions of an MNC is almost twice the probability of citation between patents from different entities *even in the same country*! Thus, not only are MNCs able to overcome the localization effect of knowledge flows, the intra-MNC knowledge flows across borders are in fact even stronger than the inter-organizational localization effect. Not surprisingly, patents that are increasingly similar in technological characteristics, like same technological category (1 of 6), same subcategory (1 of 38), same 3-digit class (1 of 410), same 9-digit primary subclass (1 of 150,000) and overlap of secondary subclasses, are also increasingly more likely to cite each other.

Table 8 extends the analysis of Table 7 to examine individual flows from domestic entities to other domestic entities (D→D), domestic entities to local subsidiaries of foreign multinationals (D→M), multinational subsidiaries to domestic entities (M→D) and multinational subsidiaries to other multinational subsidiaries (M→M). All four are found to be positive and significant, a result consistent with the earlier findings using matching (tables 2 through 5). The statistical significance of any of these coefficients indicates that the corresponding knowledge diffusion is statistically different from the reference category, which is the average cross-border inter-organizational diffusion. Compared with this reference group, D→D flows are higher by 3.0 in a million potential citations, D→M flows are higher by 3.0 in a million, M→D flows are higher by 2.1 in a million and M→M flows are higher by 4.4 in a million. Given that the average citation rate between two random patents is around 5.7 in a million, these estimates suggest that all four kinds of intra-national knowledge flows are much larger than cross-border knowledge flows across organizations. Table 8 also includes indicator variables to capture breakdown of intra-MNC flows into two cases: from a foreign subsidiary to its home base (S→H), and from its home base to its foreign subsidiary (H→S). Knowledge flows in both directions are found to be statistically quite significant, and economically several times as large as those estimated for localized inter-organizational flows of all four kinds mentioned above. Further, the comparable magnitudes of the coefficients for the two directions suggests that the extent to which MNC foreign subsidiaries contribute to knowledge of their home base is not any less important than the extent to which they draw on the knowledge of the home base. This is consistent with a view of MNCs as a "learning organization", where subsidiaries not only build upon the knowledge of the home base but also contribute to further learning (Kogut and Zander, 1993; Dunning, 1993; Nohria and Ghoshal, 1997). The finding stands in contradiction with the product life cycle model (Vernon, 1966), and much of the literature on MNCs that followed, where technical know-how was only seen as flowing one-way from the home base to the foreign subsidiaries.

The lower part of the table reports the ratio of pairs of coefficients in order to measure the relative magnitude of different kinds of diffusion.[36] The ratio $\beta_{M \to D} / \beta_{D \to M}$ is estimated to be 0.7, meaning that the coefficient for M→D flows is 30% smaller than for D→M flows. This is qualitatively consistent with our discussion of tables 4 and 5, where we had also concluded that D→M knowledge flows are stronger than M→D knowledge flows. When we try to test the result for statistical difference, we find that a test of equality of $\beta_{M \to D}$ and $\beta_{D \to M}$ is rejected at the 1% significance level.[37] Similarly, there is evidence that the M→D flows are statistically less than the D→D flows (by 30%). D→M flows, on the other hand, seem to be almost identical in strength to D→D flows. Thus, while MNCs are as good at learning from domestic organizations as domestic organizations are at learning from each other, MNCs

[36] Since the marginal effect of a variable j is therefore given by $\beta_j \Lambda'(x\beta)$, the ratio of marginal effects of two variables j and k is therefore simply β_j / β_k irrespective of the value of $\Lambda'(x\beta)$.

[37] The results in columns (1) and (2) also show all four localization effects being significant. Also, the result that the D→M effect exceeds the M→D effect also holds in both. However, the difference is not statistically significant for the latter because of high standard errors that result from the low efficiency of the procedure, as discussed earlier.

contribute somewhat less to local learning.[38] It is also interesting to note that multinational subsidiaries are particularly good at learning from each other, with the coefficient estimate being 46% higher than that for knowledge flows between domestic players, or from domestic players to MNCs. This is consistent with recent research emphasizing not just knowledge spillovers from domestic entities to MNCs, but also between MNC subsidiaries themselves, as important determinants of foreign R&D location choice by MNCs (Porter, 1990; Feinberg and Gupta, 2003).

What is the underlying mechanism for the aggregate result that knowledge flows from the host countries to the MNCs exceed those back from the MNCs to the host countries? This question is hard to answer without examining the knowledge flows in more detail. On natural possibility is to look at cross-technology differences in these patterns, since the learning-related incentives for location choice are higher for technologies where new knowledge plays an especially important role (Audretsch and Feldman, 1996). In particular, when locating abroad could lead to acquiring new knowledge, both industry laggards and leaders have an incentive to locate subsidiaries abroad. On the other hand, when the learning opportunities are small compared with the potential leakage of their existing technology, the leaders would have less incentive to locate facilities abroad.

To explore this possibility, Table 9 repeats the previous analysis for six broad technology classes defined by Hall, Jaffe and Trajtenberg (2002). I interact each of the six indicator variables discussed earlier with dummy variables for technological categories. Although this coarse technological classification surely hides heterogeneity within technological categories, some interesting patterns emerge even from this level of analysis. First, "Drugs & Medical" and "Chemical", two of the most R&D intensive sectors, show high levels of knowledge exchange among all players. This is consistent with Chung and Alcacer (2002), who suggest that these are sectors where not just the industry laggards but also the industry leaders and above-average followers actively locate advanced facilities abroad in order to keep in touch with the latest developments everywhere. On the other hand, for the "Mechanical" category, all three kinds of localized knowledge flows involving multinational

[38] In order to rule out that D→M estimate is higher than the M→D estimate due to knowledge flows from domestic universities/research labs to MNC subsidiaries, I tried a specification that included separate dummy variables for whether the D→M flows were originating from domestic firms or domestic universities/research labs. I found that

subsidiaries are much weaker than D→D flows, possibly because it is not a particularly knowledge-intensive sector. The two sectors where the aggregate result of D→M knowledge flows being significantly stronger than M→D knowledge flows still holds are "Computers & Communication" and "Electrical & Electronics". Once more, this is consistent with Chung and Alcacer's (2002) finding that Computers and Electronics are two sectors where FDI is dominated by industry laggards, with much to learn but less to contribute.

Table 10 does similar analysis for individual countries. I find evidence of strong intra-national knowledge flows all countries. However, the aggregate finding that D→M knowledge flows are stronger than M→D knowledge flows holds true only for the US, Japan and Germany. The equality cannot be rejected for France and Canada, while the trend actually reverses for the UK. One explanation for this pattern could be that the domestic firms and organizations in the US, Japan and Germany are, on an average, technologically more advanced than the average subsidiary of a foreign multinational based there, and therefore have much less to learn from the latter. Likewise, the fact that D→M knowledge flows are significantly weaker than D→D knowledge flows for the UK and Canada is consistent with an explanation that multinationals operating there (mostly from the US, Japan and Germany) are actually technologically more advanced than the domestic firms and organizations.

5. Cross-Border Effects on Parent Firm

The above analysis showed existence of *intra-national, inter-firm* knowledge flows (D→D, D→M, M→D and M→M) and *cross-border, within-firm* knowledge flows (S→H and H→S). Combining the two suggests how FDI can lead even to *cross-border, inter-firm* knowledge flows. As further evidence of *cross-border inter-firm* knowledge flows involving FDI, we can directly measure the effect of extent of FDI on cross-border patent citations between different firms (Globerman, Kokko and Sjöholm, 2000; Branstetter, 2000). To do this, I test if the probability of foreign citations by the *home base* of a firm increases with increased presence of the firm in the cited country, and with increased presence of cited assignee in the firm's home country. It is worth emphasizing that this is an extremely strong test: While detecting a positive effect indicates knowledge diffusion, a zero or negative effect

the D→M flows originating from domestic firms are actually slightly higher rather than lower than the D→M flows *does not* indicate an absence of diffusion since the diffusion operates through the indirect route mentioned above, which need not increase cross-border patent citations across firms.

I measure the presence of the citing assignee in the cited country as the logarithm of the number of its overall patents during 1986-95 originating from its foreign operations in the cited country. This can be seen as a measure of its local absorptive capacity (Cohen and Levinthal, 1989) in the country being considered. Similarly, I define the presence of the cited assignee in the citing country as the number of its overall patents during 1986-95 originating from its overseas operations in the citing country. Additional control variables are introduced for the log of worldwide patenting by the citing assignee and the cited assignee in order to ensure that the foreign presence variables do not simply pick up overall scale effects (that would arise if larger assignees systematically differ in the propensity to cite or be cited). The fixed effects and control variables from the previous section are also still used.

Since we are interested in cross-organizational flows, actual or potential self-citations are dropped. Further, since the objective is now to study only cross-border citations, all citations within the same country are also dropped. The results are reported in Table 11. As before, the dependent variable is a binary outcome that is equal to 1 if, and only if, there is an actual citation between the citing patent and the potentially cited patent. A matching-based WESML estimation procedure is again used. The negative and significant effect of the global scale of the citing assignee suggests that larger firms rely much less on external sources of knowledge, perhaps because they build on their own internal knowledge. The positive and significant effect of the global scale of the cited assignee then has a complementary explanation – innovations made by larger firms are much more frequently built upon by others than are innovations made by smaller firms.

There is a positive and significant effect for the presence of the citing assignee in the cited country. Thus increased foreign activity by a firm leads to not just higher probability of local patent citations from its subsidiary to domestic organizations in the host country (as we found in the previous two sections) but also to higher probability of cross-border citations from other domestic organizations in the host country back to the parent firm abraod. The order of magnitude of this effect is about 1.6 in a million increase in the citation probability for a 10% increase in local inventive activity abroad (recall that regressions use log of

from domestic universities/research labs.

presence, hence the percentage interpretations). Given that the average citation rate for the cross-border population of patent pairs is 5.7 in a million, this is economically non-trivial though obviously not as significant as the within-country effects we saw earlier. Similarly, there is statistically significant evidence that a 10% increase in local activity by firms from another country also increases cross-border knowledge flows from the home base of those foreign firms to the host country's domestic organizations by about 0.6 in a million. To summarize, multi-nationality seems to be associated not just with direct localized knowledge flows between foreign MNC subsidiaries and domestic firms, but also with increased cross-border knowledge flows between firms in different countries.[39]

In analysis not reported in detail here, I carried out a direct comparison with the two-country study involving patents from the US and Japan by Branstetter (2000). This involved repeating the above analysis just for the US and Japan after dropping all citations involving patents from the remaining 4 countries. Branstetter's finding that Japan's FDI in the US is a significant channel of knowledge diffusion from the indigenous US firms to the investing Japanese firms was found to hold in my analysis. However, his finding that Japan's FDI in the US is also a significant channel of knowledge diffusion from the home base of the investing Japanese firm to indigenous US firms did not find support.

I can also compare my findings with that of Globerman, Kokko and Sjöholm (2000), who use citation data from the Swedish patent and employment data to measure inward and outward FDI for Sweden. My findings are similar to theirs – firms in a country learn more from outward FDI than from inward FDI. Given that the authors reach this conclusion using data from the Swedish patent office rather than the US Patent Office, their study provides additional evidence that my own results are not merely a result of biases arising from use of US patent data.

[39] When analysis analogous to Table 11 was carried out separately for each of the six countries, increased presence of citing assignee in cited country is shown to have positive and significant effect in five of the six countries: US, Japan, France (at 10% significance level), UK and Canada. Similarly, increased presence of potentially cited assignee in the citing country is shown to have a positive and significant effect in 2 of the countries: Japan and Canada (at 10% level). As already pointed out, however, the fact that the finding is not robust across countries might simply be the result of the test carried out being quite stringent: since inter-firm knowledge flows involving MNCs are mainly through localized knowledge exchange with their local subsidiaries, the indirect effect on cross-border citations might or might not be significant.

6. Appropriateness of USPTO Data for Measuring Knowledge Flows

As discussed earlier in the paper, all my regressions include country fixed effects to control for the fact that inventors based in different countries might systematically differ in their propensity to cite USPTO patents. However, this does not resolve a related concern that MNC subsidiaries and domestic organizations *even within the same country* might differ in the extent to which they choose not to cite USPTO patents, and instead cite the same innovations as registered with a different patent office like European Patent Office (EPO).[40] Since such data points would be absent from the NBER patent citations data (which only includes citations to US patents), this could lead to biases in the measured knowledge flows. To look into this possibility, I looked at citation patterns for a random sample of 1,612 USPTO patents from 1995, about half being from patents by domestic organizations and the other half being from patents by MNC subsidiaries. In particular, I manually inspected citations made to the EPO patents from this sample. For each patent, I identified which of the cited EPO patents had an equivalent USPTO patent that could have instead been cited, and therefore was "missing" from the NBER database of patent citations. The mapping from EPO to USPTO patents was done using the "OECD Triadic Patent Families" database, which has information on patents filed for the same innovation at both USPTO and EPO.

The results are summarized in Table 12. The mean number of citations to USPTO patents by patents in the above sample was 5.85 citations. In contrast, there were an average of 1.13 citations to EPO patents. However, about two-thirds of the EPO citations did not have a corresponding USPTO patent. The mean number of citations to EPO patents with equivalent USPTO patents, which measures the average number of "missing citations" as described above, was 0.32 for a random patent. In other words, most of the citations made by USPTO patents are to other USPTO patents rather than EPO patents, and hence do show up in my data. As Table 12 shows, the average number of "missing citations" per patent from MNC subsidiaries (0.43) is a little higher than that per patent from domestic organizations (0.22). This holds both in the sub-sample of patents originating in the US (0.39 for MNC subsidiaries and 0.24 for domestic organizations), and for those that originate elsewhere (0.46 for MNC

[40] USPTO patents provide patent protection only in the US. For protection in Europe as well, a patent needs to be separately applied for in Europe. EPO makes it possible to get patent protection in several European countries through a centralized application process, with the filing fees determined by the number of countries in which patent protection is sought.

subsidiaries and 0.21 for domestic organizations). In either sub-sample, the missing citation bias therefore is in the direction of underestimating the extent of localized knowledge diffusion to MNCs more than to domestic organizations. In other words, if we could correct for this bias in the previous analysis, it would only strengthen the previous result that D→M flows exceed M→D flows.

7. Conclusion

This paper studies the role of multinational firms (MNCs) in international knowledge diffusion. This is done by measuring micro-level knowledge flows using citation data from around half a million US Patent Office (USPTO) patents from 1986 to 1995, covering innovations from six countries and all manufacturing sectors. Cross-border knowledge flows within an MNC are shown to be several times stronger than those not within the same MNC, hence presenting econometric evidence for the conventional wisdom that MNCs are particularly good at cross-border knowledge transfer. I find significant localized knowledge flows from domestic firms and organizations to foreign MNC subsidiaries, and also from foreign MNC subsidiaries to domestic entities. However, while the former is about as strong as domestic-to-domestic knowledge diffusion, the latter is found to be weaker in the aggregate analysis. Thus, MNCs seem to learn more from their host countries than the host countries learn from MNCs. I also find direct evidence of multi-nationality being a channel for inter-organizational knowledge diffusion across national boundaries, with the probability of knowledge flow to an MNC's home base being an increasing function of the extent of its innovative activity in a foreign country.

This paper also explores how the aggregate findings vary across countries and industries. In particular, the result that knowledge flows from domestic organizations to MNCs exceed those from MNCs back to domestic organizations only holds for three of the six countries (US, Japan and Germany), and two of the six broad technological categories ("Computers & Communications" and "Electrical & Electronics"). As discussed earlier, this is consistent with the view that knowledge spillovers from foreign MNC subsidiaries to domestic players are stronger in the countries and industries where these subsidiaries are technologically more advanced. This illustrates two more general issues that need to be addressed more explicitly in international economics: First, the overseas location choice for multinational subsidiaries may depend not just on the demand or technology characteristics already emphasized in research, but also on learning opportunities for MNCs in the foreign locations. Second, since the learning opportunities vary depending on an MNC's existing technological capability, the determinants of learning-motivated decisions to locate abroad would depend on characteristics not only the home country but also of the firm itself. Third, not just the extent of foreign investment but also its quality, relative to domestic investment in the host country, needs to be taken into account in studying the relative extent of knowledge spillovers to and from MNCs.

Since the types of knowledge flows captured by patent citation data are more relevant for technologically advanced countries, this paper has focused only on six developed countries. The results presented here might not be a good representation of developing countries. In particular, since the domestic organizations are often technologically far behind MNCs in developing countries, the learning effect might be much weaker for the MNCs and much stronger for the domestic organizations instead. In addition, such learning probably leads less often to radical innovation and more often to direct adoption or incremental improvement of existing technologies. In fact, different kinds of MNC activity, e.g., R&D laboratories versus manufacturing plants, might have very different implications for learning not just in the developing but also in the developed countries. It is therefore no surprise that the literature on FDI finds mixed results on knowledge spillovers - there is probably no blanket "truth" about knowledge diffusion involving MNCs. The extent and direction of diffusion is surely sensitive to the type of FDI, the characteristics of the MNCs, and the existing level of technology of the home as well as host country. Thus, future research should focus much less on the *whether* and much more on the *when* of knowledge spillovers. That is the only way to arrive at useful implications for MNC who are managers making decisions on locating facilities abroad, and for policymakers who are trying to encourage knowledge spillovers from MNCs to the domestic economy.

An issue that most papers on knowledge spillovers, including this one, leave open is: what are the *mechanisms* behind knowledge diffusion? For example, *why* do knowledge spillovers tend to be geographically localized? *How* do MNCs overcome this geographic constraint on knowledge diffusion and manage to transmit knowledge across large distances? In a related paper (Singh, 2003), I present a framework for explaining patterns of knowledge diffusion based on the premise that individual inventors are carriers of tacit knowledge. In particular, I show that inventor mobility and social networks shaped by the history of collaborations of inventors are useful in explaining some of the observed patterns of knowledge diffusion. In future research, I would like to continue this exploration into the mechanisms of knowledge diffusion, focusing specifically on international knowledge diffusion resulting from cross-border labor mobility as well as social networks that cross national borders, both within and outside multinational firms.

8. References

Aitken, B., and A. Harrison (1999), "Do Domestic Firms Benefit from Foreign Investment? Evidence from Venezuela" *American Economic Review* 89: 605-618

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

Arrow, K. (1969), "Classificatory notes on the production and transmission of technological knowledge" *American Economic Review Paper and Proceedings* 59: 29-35.

Amemiya, T. (1985) *Advanced Econometrics*. Harvard University Press, Cambridge.

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.

Baldwin, R., H. Braconier, and R. Forslid (1999), "Multinationals, endogenous growth and technological spillovers: Theory and evidence" CEPR Discussion Paper #2155.

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

Birkinshaw, J.M. (1997), "Entrepreneurship in Multinational Corporations: The Characteristics of Subsidiary Initiatives." *Strategic Management Journal*, 18(3): 207-229.

Blomström, M. and A. Kokko (1998), "Multinational Corporations and Spillovers," *Journal of Economic Surveys* 12(3): 247-277.

Branstetter, L. (2000), "Is foreign direct investment a channel of knowledge spillovers? Evidence from Japan's FDI in the United States" NBER Working Paper # 8015.

Branstetter, L. (2001), "Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the U.S. and Japan" *Journal of International Economics* 53: 53-79. 53

Cantwell, J. (1989) *Technological Corporations and Innovatory Activities*. Oxford; Basil Blackwell.

Caves, R.E. (1974), "Multinational firms, competition and productivity in host-country markets" *Economica*, 41: 176-193.

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.

Coe, D.T., and E. Helpman (1995), "International R&D Spillovers" *European Economic Review* 39: 859-887.

Cohen, W., and D. Levinthal (1989), "Innovation and learning: The two faces of R&D", *Economic Journal* 99: 569-596.

Coslett, S.R. (1981), "Maximum Likelihood Estimator for Choice-based Samples." *Econometrica* 49: 1289-1316.

Dalton, D.H. and M.G. Shapiro (1995). *Globalizing Industrial Research & Development*. Office of Technology Policy, US Dept of Commerce.

Duget, E. and M. MacGarvie (2002), "How Well Do Patent Citations Measure Knowledge Spillovers?" Working paper.

Dunning, J.H. (1993). *Multinational Enterprises and the Global Economy.* Addison-Wesley.

Eaton, J., and S. Kortum (1996), "Trade in Ideas: Patenting and Productivity in the OECD", *Journal of International Economics* 40: 251-278.

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

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

Girma, S., and K. Wakelin (2001), "Regional underdevelopment: Is FDI the solution? A semi-parametric analysis", GEP Research Paper 2001/11, University of Nottingham, U.K.

Globerman, S., A. Kokko, and F. Sjöholm (2000), "International Technology Diffusion: Evidence from Swedish Patent Data", *Kyklos* 53: 17-38. 55

Griliches, Z. (1990), "Patent statistics as economic indicators: A survey" *Journal of Economic Literature* 28: 1661-1797.

Greene, W. (2003) *Econometric Analysis*. Prentice Hall, 5th Edition.

Hall, B.H., A.B. Jaffe, and M. Trajtenberg (2002), "The NBER patent citations data file: lessons, insights and methodological tools" In Jaffe and M. Trajtenberg (2002).

Håkanson, L. and R. Nobel (1993), "Determinants of foreign R&D in Swedish multinationals" *Research Policy* 22(5-6): 397-411.

Hanson, G.H. (2001), "Should countries promote foreign direct investment?" G-24 Discussion Paper Series, UNCTAD.

Haskel, J., S. Pereira, and M. Slaughter (2001), "Does inward foreign direct investment boost the productivity of domestic firms?" Presented at the NBER Summer Institute, August.

Hedlund, G. (1986) 'The Hypermodern MNC - A Heterarchy?' *Human Resource Management*, 25: 9-35.

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

Imbens, G (1992), "An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-based Sampling." *Econometrica* 60: 1187-1214.

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., M. Trajtenberg and M.S. Fogarty (2000), "Knowledge spillovers and patent citations: evidence from a survey of inventors" *American Economic Association Papers and Proceedings* 90(2).

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

Keller, W. (1998), "Are international R&D spillovers trade related? Analyzing spillovers among randomly matched trade partners", *European Economic Review* 42: 1469-1481.

Keller, W. (2002), "Geographic localization of international technology diffusion", *American Economic Review* 92(1).

King, G. and L. Zeng (2001), "Logistic Regression in Rare Events Data", *Political Analysis* 9(2): 137-163

Kogut, B. and S.J. Chang (1991), "Technological Capabilities and Japanese Foreign Direct Investment in the United States" *The Review of Economics and Statistics* 73 (3): 401-413.

Kogut, B. and U. Zander (1993), "Knowledge of the Firm and the Evolutionary Theory of the Multinational Corporation." *Journal of International Business Studies*. 24(4), pp. 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* 28: 179-193.

Lancaster, Tony and G. Imbens (1996),"Efficient Estimation and Stratified Sampling." *Journal of Econometrics* 74: 289-318.

Lerner, J. (2002), "150 Years of Patent Protection" *American Economic Review* 92 (2).

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.

Lichtenberg, F., and B. van Pottelsberghe de la Potterie (1998), "International R&D spillovers: A Comment", *European Economic Review* 42 (8): 1483-1491.

Manski, C.F. and S.R. Lerman (1977), "The Estimation of Choice Probabilities from Choice Based Samples," *Econometrica* 45(8): 1977-88.

MacGarvie, M. (2003), "Do firms learn from international trade? Evidence from patent citations and micro data," Mimeo.

Moran, T.H. (2001) *Parental Supervision: The New Paradigm for Foreign Direct Investment and Development*, Institute of International Economics.

Nohria, N. and S. Ghoshal (1997). *The Differentiated network: Organizing Multinational Corporations for Value Creation.* Jossey-Bass Publishers, San Francisco.

Peri, G. (2002), "Knowledge Flows and Innovation," mimeo.

Porter, M.E. (1990) *The Competitive Advantage of Nations*. Free Press.

Rodrik, D. (1999). *The New Global Economy And Developing Countries: Making Openness Work*. John Hopkins University Press for the Overseas Development Council.

Singh, J. (2003), "Inventor Networks and Inter-Firm Knowledge Diffusion," mimeo.

Sorenson, O. and L. Fleming (2001) "Science and the Diffusion of Knowledge." Working paper 02-095, Harvard Business School.

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

Thompson, P. and M. Fox-Kean (2003), "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment." Working Paper.

Vernon, R. (1966), "International Investment and International Trade in the Product Cycle." *Quarterly Journal of Economics* 80: 190-207.

Xu, B. (2000), "Multinational Enterprises, Technology Diffusion, and Host Country Productivity Growth", *Journal of Development Economics,* 62: 477-493.

Appendix A: Cleaning up patent assignee data

I undertook the following extensive data cleaning exercise for the patent data. First, I used Compustat-based CUSIP numbers (from year 1989) included in NBER database to make sure that assignees that have the same CUSIP numbers are matched to the same parent. The number of assignees with available CUSIP identification numbers was around 4,600, and the number of unique parent CUSIPs to which they could be matched was around 2,500. I then used Stopford's *Directory of Multinationals* from 1992 to manually associate subsidiaries of 428 of the largest MNCs to their respective parents. I was able to associate 2,700 assignees (which do not have CUSIP ids in the database) with about 200 corporate parents, many of which were non-US MNCs. In addition, I reexamined the corporate ownership of the assignees with CUSIP numbers examined above, and found the need to correct the parent assignment for about 100 of them. Next, using a combination of inspection of USPTO assignee codes, keyword search and verification using the Internet, I identified about 400 government-affiliated bodies, 550 research institutes and 450 universities worldwide. Finally, the ultimate ownership of another 1,000 major patent assignees was manually checked using a combination of *Who Owns Who* directories for the year 1991 and Internet search. As Table 1 shows, the total of around 10,000 assignees inspected in the above steps accounts for 555,643 patents for 1986-1995, which is about 73% of all assigned patents from the six countries studies.

Appendix B: Combining WESML with Matching

The choice-based WESML procedure (Manski and Lerman, 1977) can be extended to a sample obtained by matching, i.e., by taking all actual citations (denoted by y=1) and matching each of these with *k* "control citations" (denoted by y=0) along a dimension *z* (e.g., the "cells" indexed by the vector combination of the citing technological class and cited technological class). Without loss generality, denote the values *z* can take as 1, 2, …, *T*. For a matching-based sampling design, it is easier to think of not just y but (z, y) as the dependent variable. In forming the likelihood function, I will use the result that

$$\begin{aligned}\Pr(z = z_i \ and \ y = j \mid x = x_i) &= \Pr(z = z_i \mid x_i)\Pr(y = j \mid z = z_i \ and \ x = x_i) \\ &= \Pr(z = z_i \mid x_i)\Pr(y = j \mid x = x_i)\end{aligned}$$

The second equality assumes that the vector **x** includes all information about *z* that affects citation outcome *y*, i.e., **x** is a sufficient statistic for z.[41] The log likelihood function for estimation using an exogenous random sample of size n would therefore be

$$\ln L = \sum_{i=1}^{n} \ln\left[\Pr(z = z_i \ \ and \ \ y = y_i \mid x_i)\right]$$

$$= \sum_{i=1}^{n} \left\{ y_i \ln\left[\Pr(z = z_i \mid x_i)\Lambda(x_i\beta)\right] + (1 - y_i)\ln\left[\Pr(z = z_i \mid x_i)\left(1 - \Lambda(x_i\beta)\right)\right]\right\}$$

This forms the basis of driving the pseudo-likelihood function for choice-based sampling. The procedure involves weighting each of the original log likelihood function term by the inverse of the probability that the corresponding population element will be included in the sample. To derive these weights, denote the number of elements with $z = t$ and y=j as n_{tj} for the sample and N_{tj} for the population (t is between 1 and *T*, and j is either 0 or 1). The matched sampling implies that we pick, from each cell, all elements with y=1 and k times as many elements with y=0. In other words, $n_{t1} = N_{t1}$ and $n_{t0} = kN_{t1}$. Also, since we know N_{tj}, the probability p_{tj} of a population element with z = t and y = j getting selected in our sample is easily calculated as $p_{t1}= n_{t1}/N_{t1}=1$ and $p_{t0}= n_{t0}/N_{t0} = kN_{t1}/N_{t0}$ for all values of t. Denoting $w_{tj} = 1/p_{tj}$, the weighted likelihood function for choice-based sampling is the given by

$$\ln L_w = \sum_{i=1}^{n} \left\{ y_i w_{z_i 1} \ln\left[\Pr(z = z_i \mid x_i)\Lambda(x_i\beta)\right] + (1 - y_i) w_{z_i 0} \ln\left[\Pr(z = z_i \mid x_i)\left(1 - \Lambda(x_i\beta)\right)\right]\right\}$$

$$= C - \sum_{i=1}^{n} w_i \ln\left(1 + e^{(1-2y_i)X\beta}\right)$$

where $\quad w_i = y_i w_{z_i 1} + (1 - y_i) w_{z_i 0} \quad$ and $\quad C = \sum_{i=1}^{n} w_i \ln\left[\Pr(z = z_i \mid x_i)\right]$

Since C is independent of $\boldsymbol{\beta}$, it can be ignored in the maximum likelihood procedure. Thus we can again use weighted logit estimation, where the weights of the observations are now given by w_i. Unlike the simple WESML with random sampling from the y=0 observations, the weights now depend not just on the value of y but also on the cell that the observations falls into.

---

[41] In the most conservative view, this assumption means that one needs to include dummy variables for all possible values of z (i.e., all technological class combinations) in the vector **x**. My actual regressions are less conservative and assume that the effects of technological classes on citation probabilities are completely captured by dummy variables for broad technological categories and three dummy variables for whether the citing and cited patents are in the same broad category, subcategory or technological class.

Table 1: Countries and patents included in the sample

This table lists the patenting activity arising from the 6 countries included in the analysis in this paper. The country in which a patent originates is determined using the address of the first inventor. When the assignee country is the same as the inventor country, the patent is classified as being owned by a "domestic firm or organization". When the assignee country is different from the inventor country, the patent is classified as being owned by a "multinational subsidiary". The first column below lists all USPTO patents for a given country for 1986-95, while the second column only includes patents that are owned by a firm or organization rather than being owned by individuals. The third and fourth columns (as well as all the tables that follow) are based only on the subset of these patents for which the assignee information has been inspected to be correctly pointing to the parent firm or organization rather than possibly just a division or a subsidiary of another firm. Application year is always used for classifying patents.

Country	Total patents 1986-95 in NBER database	Total number of assigned patents	Assigned patents with clean parent information	Fraction of patents from multinational subsidiaries
	(1)	(2)	(3)	(4)
United States	546,824	418,045	287,787	8.5%
Japan	217,313	212,427	183,870	2.1%
Germany	74,041	67,154	45,869	19.5%
France	29,791	27,120	17,289	20.4%
United Kingdom	26,631	23,968	15,131	40.3%
Canada	20,700	13,015	5,697	50.0%
Subtotal 6 countries	915,300	761,729	555,643	9.0%
Other countries	94,924	73,115	38,402	27.3%
Total worldwide	1,010,224	834,844	594,045	10.2%

Table 2: Matching-based analysis of knowledge diffusion between domestic organizations (D→D)

Column (1) gives the total number of citations made *by* domestic organizations, and columns (2) and (3) respectively give the number and fraction of these that were made *to* patents by domestic organizations. Columns (4) and (5) report the same analysis for "control citations" obtained by replacing each original cited patent by a "control patent" matched by 9-digit patent technological subclass, and having application date within one year of the original. All citations for which a control could not be found, or for which either the original or the control involved a self-cite have been excluded from the analysis. Column (6) reports the difference of the proportions from columns (3) and (5), column (7) calculates the t-statistic for testing the equality of the two proportions, and column (8) gives the ratio of the two proportions (a measure of intensity of localized D→D knowledge diffusion).

		Actual Citations		Control Citations		Comparison		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	**Total citations by domestic**	**Citations by domestic to domestic**	**%Citations by domestic to domestic**	**Citations by domestic to domestic**	**%Citations by domestic to domestic**	**(3) - (5)**	**t-ratio**	**(3)/(5)**
United States	430,262	243,600	56.62%	200,104	46.51%	10.11%	94.3	1.22
Japan	245,441	131,005	53.38%	108,085	44.04%	9.34%	65.7	1.21
Germany	27,326	2,989	10.94%	2,156	7.89%	3.05%	12.2	1.39
France	12,727	528	4.15%	291	2.29%	1.86%	8.4	1.81
United Kingdom	7,895	450	5.70%	207	2.62%	3.08%	9.7	2.17
Canada	3,536	100	2.83%	40	1.13%	1.70%	5.1	2.50
Total	**727,187**	**378,672**	**52.07%**	**310,883**	**42.75%**	**9.32%**	**113.1**	**1.22**

Table 3: Knowledge diffusion from domestic organizations to MNC subsidiaries (D→M)

Column (1) gives the total number of citations made *by* subsidiaries of foreign multinationals, and columns (2) and (3) respectively give the number and fraction of these that were made *to* patents by domestic organizations. Columns (4) and (5) report the same analysis for "control citations" obtained by replacing each original cited patent by a "control patent" matched by 9-digit patent technological subclass, and having application date within one year of the original. All citations for which a control could not be found, or for which either the original or the control involved a self-cite have been excluded. Column (6) reports the difference of the proportions from columns (3) and (5), column (7) calculates the t-statistic for testing the equality of the two proportions, and column (8) gives the ratio of the two proportions (a measure of intensity of localized D→M knowledge diffusion).

		Actual Citations		Control Citations		Comparison		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	**Total citations by mult sub**	**Citations by mult sub to domestic**	**%Citations by mult sub to domestic**	**Citations by mult sub to domestic**	**%Citations by mult sub to domestic**	**(3) - (5)**	**t-ratio**	**(3)/(5)**
United States	41,272	22,590	54.73%	18,799	45.55%	9.19%	26.5	1.20
Japan	5,156	2,464	47.79%	2,083	40.40%	7.39%	7.6	1.18
Germany	10,841	1,302	12.01%	985	9.09%	2.92%	7.0	1.32
France	3,856	166	4.30%	114	2.96%	1.35%	3.2	1.46
United Kingdom	9,689	220	2.27%	274	2.83%	-0.56%	-2.5	0.80
Canada	3,457	38	1.10%	25	0.72%	0.38%	1.6	1.52
Total	**74,271**	**26,780**	**36.06%**	**22,280**	**30.00%**	**6.06%**	**24.9**	**1.20**

Table 4: Knowledge diffusion from MNC subsidiaries to domestic organizations (M→D)

Column (1) gives the total number of citations made *by* domestic organizations, and columns (2) and (3) respectively give the number and fraction of these that were made *to* patents by local subsidiaries of foreign multinationals. Columns (4) and (5) report the same analysis for "control citations" obtained by replacing each original cited patent by a "control patent" matched by 9-digit patent technological subclass, and having application date within one year. All citations for which a control could not be found, or for which either the original or the control involved a self-cite have been excluded from the analysis. Column (6) reports the difference of the proportions from columns (3) and (5), column (7) calculates the t-statistic for testing the equality of the two proportions, and column (8) gives the ratio of the two proportions (a measure of intensity of localized M→D knowledge diffusion).

		Actual Citations		Control Citations		Comparison		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	**Total citations by domestic**	**Citations by domestic to mult sub**	**%Citations by domestic to mult sub**	**Citations by domestic to mult sub**	**%Citations by domestic to mult sub**	**(3) - (5)**	**t-ratio**	**(3)/(5)**
United States	430,262	17,010	3.95%	15,136	3.52%	0.44%	10.7	1.12
Japan	245,441	2,082	0.85%	1,879	0.77%	0.08%	3.2	1.11
Germany	27,326	658	2.41%	542	1.98%	0.42%	3.4	1.21
France	12,727	124	0.97%	101	0.79%	0.18%	1.5	1.23
United Kingdom	7,895	197	2.50%	149	1.89%	0.61%	2.6	1.32
Canada	3,536	32	0.90%	15	0.42%	0.48%	2.5	2.13
Total	**727,187**	**20,103**	**2.76%**	**17,822**	**2.45%**	**0.31%**	**11.9**	**1.13**

Table 5: Knowledge diffusion between MNC subsidiaries (M→M)

Column (1) gives the total number of citations made *by* multinational subsidiaries, and columns (2) and (3) respectively give the number and fraction of these that were made *to* patents by local subsidiaries of foreign multinationals. Columns (4) and (5) report the same analysis for "control citations" obtained by replacing each original cited patent by a "control patent" matched by 9-digit patent technological subclass, and having application date within one year. All citations for which a control could not be found, or for which either the original or the control involved a self-cite have been excluded from the analysis. Column (6) reports the difference of the proportions from columns (3) and (5), column (7) calculates the t-statistic for testing the equality of the two proportions, and column (8) gives the ratio of the two proportions (a measure of intensity of localized M→M knowledge diffusion).

		Actual Citations		Control Citations		Comparison		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	**Total citations by mult sub**	**Citations by mult sub to mult sub**	**%Citations by mult sub to mult sub**	**Citations by mult sub to mult sub**	**%Citations by mult sub to mult sub**	**(3) - (5)**	**t-ratio**	**(3)/(5)**
United States	41,272	1,984	4.81%	1,590	3.85%	0.95%	6.7	1.25
Japan	5,156	98	1.90%	66	1.28%	0.62%	2.5	1.48
Germany	10,841	241	2.22%	171	1.58%	0.65%	3.5	1.41
France	3,856	68	1.76%	28	0.73%	1.04%	4.1	2.43
United Kingdom	9,689	215	2.22%	258	2.66%	-0.44%	-2.0	0.83
Canada	3,457	21	0.61%	16	0.46%	0.14%	0.8	1.31
Total	**74,271**	**2,627**	**3.54%**	**2,129**	**2.87%**	**0.67%**	**7.3**	**1.23**

Table 6: Summary of variables used for regressions analysis

Same tech category	Indicator variable that is 1 if both the citing and the potentially cited patent belong to the same broad industry category (one of 6) as defined in the NBER database
Same tech subcategory	Indicator variable that is 1 if both the citing and the potentially cited patent belong to the same broad technical subcategory (one of 36) as defined in the NBER database
Same primary tech class	Indicator variable that is 1 if both the citing and the potentially cited patent belong to the same 3-digit primary technology class (one of about 450) as defined in the US Patent classification system
Same primary subclass	Indicator variable that is 1 if both the citing and the potentially cited patent belong to the same 9-digit primary technology subclass (one of about 150,000) as defined in the US Patent classification system
Secondary subclass overlap	Indicator variable that is 1 if at least one of the secondary 9-digit subclasses of one patent is the same as a primary or secondary subclass of the other patent in the dyad
Within same country	Indicator variable that is 1 if the citing and cited patents originate from inventors located in the same country
Within same MNC	Indicator variable that is 1 if the citing and cited patents are from two divisions (located in different countries) of the same MNC
D→D	Indicator variable that is 1 if both the citing and potentially cited patent belong to the same country, with assignees for both being domestic players in the country
D→M	Indicator variable that is 1 if both the citing and potentially cited patent belong to the same country, with assignee for the former being a local subsidiary of a foreign multinational and for the latter being a domestic player
M→D	Indicator variable that is 1 if both the citing and potentially cited patent belong to the same country, with assignee for the former being a domestic player and for the latter being a local subsidiary of a foreign multinational
M→M	Indicator variable that is 1 if both the citing and potentially cited patent belong to the same country, with assignees for both local subsidiaries of foreign multinationals
S→H	Indicator variable that is 1 if citing patent is from the home base of an MNC and the cited patent is from a foreign subsidiary (located abroad) of the same MNC
H→S	Indicator variable that is 1 if citing patent is from the local subsidiary of a foreign MNC and the cited patent is from the home base (located abroad) of the same MNC
Presence of citing assignee in cited country	Log(1 + number of patents that originate in the same country as the potentially cited patent and are assigned to the citing entity)
Presence of cited assignee in citing country	Log(1 + number of patents that originate in the same country as the citing patent and are assigned to the potentially cited entity)
Scale of citing assignee	Log(number of worldwide patents for 1980-99 that are assigned to the citing entity)
Scale of cited assignee	Log(number of worldwide patents for 1980-99 that are assigned to the cited entity)

Table 7: Regression approach for calculating knowledge diffusion

The sample here consists of actual or potential patent citations. The dependent variable is 1 if and only if there is an actual citation between the two patents, and 0 otherwise. The regression model used is a weighted logit based on choice-based sampling. The weights for term of the likelihood function are calculated as the inverse of the ex ante probability of its inclusion in the sample. This tables shows that knowledge diffusion is particularly high within a same country and within a same MNC, even after carefully controlling for technological relatedness of patents. It also shows that inadequate controls for technology (e.g. just at the 3-digit class level, as used in most existing studies and reproduced in column 2) can bias the knowledge spillover results. Column (4), where technological relatedness of patents is captured at the 9-digit technology level, and considers both the primary and additional technological classification, is the specification of choice. The rest of the rest of the paper therefore uses this specification, though the coefficients on the technology variables are not shown in the remaining tables in order to enhance readability.

	(1)	(2)	(3)	(4)
Within same country	0.672**	0.578**	0.553**	0.520**
	(0.009)	(0.005)	(0.006)	(0.009)
	[3.83]	[3.29]	[3.15]	[2.96]
Within same MNC	3.291**	2.110**	1.969**	1.825**
	(0.110)	(0.026)	(0.045)	(0.050)
	[18.76]	[12.03]	[11.22]	[10.40]
Same tech category		1.148**	1.144**	1.108**
		(0.011)	(0.011)	(0.012)
		[6.54]	[6.52]	[6.32]
Same tech subcategory		1.246**	1.256**	1.218**
		(0.014)	(0.014)	(0.015)
		[7.10]	[7.16]	[6.94]
Same primary tech class		3.243**	2.963**	1.930**
		(0.011)	(0.011)	(0.015)
		[18.49]	[16.89]	[11.00]
Same primary subclass			3.539**	2.282**
			(0.014)	(0.028)
			[20.17]	[13.01]
Secondary subclass overlap				4.111**
				(0.012)
				[23.43]
Number of observations	5,577,206	5,577,206	5,577,206	5,577,206

Robust standard errors in parentheses, with clustering on citing patent
Marginal effects in square brackets after multiplication with 1,000,000
Fixed effects used for technological category of citing patent, country of citing patent, citing patent year and time lag between patents
** significant at 1%; * significant at 5%

Table 8: Knowledge diffusion to and from multinational subsidiaries

This table investigates the findings from column (4) of Table 7 at a more detailed level. The "within same country" category is now broken down into four sub-categories to capture all four possible directions of knowledge flows among multinational subsidiaries and domestic players. The "within same MNC" category for intra-MNC knowledge flows across borders is now broken down into knowledge flows from subsidiaries to home base and vice versa. The ratios included at the end of the table also indicate the results for a statistically test of equality of the estimated coefficients. Among intra-national flows, domestic-to-multinational flows are found to be about as strong as domestic-to-domestic flows, while multinational-to-domestic flows are weaker and multinational-to-multinational-flows are stronger. Among intra-MNC flows across borders, subsidiary-to-home base flows and home base-to-subsidiary flows are found to be of comparable magnitude.

Within same country	
D→D	0.525** (0.010) [2.99]
D→M	0.521** (0.032) [2.97]
M→D	0.366** (0.030) [2.09]
M→M	0.768** (0.096) [4.38]
Within same MNC	
S→H	1.796** (0.080) [10.24]
H→S	1.848** (0.061) [10.53]
Observations	5,577,206
$\beta_{D\to M} / \beta_{D\to D}$	0.99
$\beta_{M\to D} / \beta_{D\to D}$	0.70**
$\beta_{M\to M} / \beta_{D\to D}$	1.46**
$\beta_{M\to D} / \beta_{D\to M}$	0.70**
$\beta_{H\to S} / \beta_{S\to H}$	1.03

Robust standard errors in parentheses, with clustering on citing patent
Marginal effects in square brackets after multiplication with 1,000,000
Controls for technological similarity of citing and cited patent included in regression, but not shown here to enhance readability
Fixed effects used for technological category of citing patent, country of citing patent, citing patent year and time lag between patents
** significant at 1%; * significant at 5% (In case of ratios, whether statistically different from 1 is tested)

Table 9: Estimation of intra-national knowledge diffusion, by technological category

This table repeats analysis of intra-national knowledge diffusion after interacting dummy variables of broad technological category with the coefficients of interest. This helps study how the patterns of knowledge diffusion vary across industries.

	Technological category of citing patent					
	Chemical	**Computers & Communications**	**Drugs & Medical**	**Electrical & Electronic**	**Mechanical**	**Other**
Within same country						
D→D	0.473**	0.498**	0.702**	0.485**	0.575**	0.631**
	(0.023)	(0.016)	(0.053)	(0.019)	(0.019)	(0.043)
D→M	0.512**	0.623**	0.700**	0.439**	0.270**	0.667**
	(0.059)	(0.053)	(0.165)	(0.071)	(0.082)	(0.100)
M→D	0.491**	0.237**	0.665**	0.150*	0.289**	0.799**
	(0.059)	(0.060)	(0.091)	(0.076)	(0.066)	(0.110)
M→M	0.548**	0.902**	1.483**	0.682*	0.209	1.763**
	(0.189)	(0.147)	(0.184)	(0.271)	(0.200)	(0.214)
Within same MNC						
S→H	1.691**	1.563**	2.221**	1.571**	2.378**	1.891**
	(0.168)	(0.124)	(0.285)	(0.179)	(0.188)	(0.371)
H→S	1.922**	1.539**	2.208**	1.679**	1.899**	2.504**
	(0.124)	(0.101)	(0.173)	(0.134)	(0.167)	(0.354)
Category fixed effect	-	1.027**	-0.714**	0.585**	0.618**	-0.382**
		(0.019)	(0.037)	(0.019)	(0.018)	(0.030)
$\beta_{D \to M} / \beta_{D \to D}$	1.08	1.25*	1.00	0.91	0.47**	1.06
$\beta_{M \to D} / \beta_{D \to D}$	1.04	0.48**	0.95	0.31**	0.50**	1.27
$\beta_{M \to M} / \beta_{D \to D}$	1.16	1.81**	2.11**	1.41	0.36	2.79**
$\beta_{M \to D} / \beta_{D \to M}$	0.96	0.38**	0.95	0.34**	1.07	1.20
$\beta_{H \to S} / \beta_{S \to H}$	1.14	0.98	0.99	1.07	0.80	1.32

Robust standard errors in parentheses, with clustering on citing patent

Controls for technological similarity of citing and cited patent included in regression, but not shown here to enhance readability

Fixed effects used for technological category of citing patent, country of citing patent, citing patent year and time lag between patents

** significant at 1%; * significant at 5% (In case of ratios, whether statistically different from 1 is tested)

Table 10: Estimation of intra-national knowledge diffusion, by country

This table repeats analysis of intra-national knowledge diffusion after interacting dummy variables of individual countries with the coefficients of interest. This helps study how the patterns of knowledge diffusion vary across countries.

	Country of origin of citing patent					
	US	**Japan**	**Germany**	**France**	**UK**	**Canada**
Within same country						
D→D	0.517**	0.535**	0.503**	0.526**	0.688**	1.406**
	(0.013)	(0.016)	(0.042)	(0.089)	(0.141)	(0.173)
D→M	0.491**	0.579**	0.941**	0.700**	0.281*	0.865**
	(0.037)	(0.081)	(0.114)	(0.148)	(0.109)	(0.213)
M→D	0.371**	0.255*	0.461**	0.719**	0.670**	1.015**
	(0.032)	(0.103)	(0.082)	(0.149)	(0.143)	(0.245)
M→M	0.695**	1.357**	0.633**	1.738**	0.934**	1.061**
	(0.120)	(0.354)	(0.235)	(0.338)	(0.167)	(0.309)
Within same MNC						
S→H	1.925**	1.771**	1.153**	1.357**	1.920**	2.383**
	(0.107)	(0.212)	(0.204)	(0.192)	(0.211)	(0.292)
H→S	1.607**	2.097**	2.203**	1.964**	1.644**	2.177**
	(0.115)	(0.251)	(0.145)	(0.120)	(0.095)	(0.100)
Country fixed effect	-	-0.384**	-0.319**	-0.248**	-0.064	-0.022
		(0.014)	(0.021)	(0.018)	(0.038)	(0.028)
$\beta_{D\to M} / \beta_{D\to D}$	0.95	1.08	1.87**	1.33	0.41*	0.62*
$\beta_{M\to D} / \beta_{D\to D}$	0.72**	0.48**	0.92	1.37	0.97	0.72
$\beta_{M\to M} / \beta_{D\to D}$	1.34	2.54*	1.26	3.30**	1.36	0.75
$\beta_{M\to D} / \beta_{D\to M}$	0.76**	0.44**	0.49**	1.03	2.38*	1.17
$\beta_{H\to S} / \beta_{S\to H}$	0.83*	1.18	1.91**	1.45**	0.86	0.91

Robust standard errors in parentheses, with clustering on citing patent

Controls for technological similarity of citing and cited patent included in regression, but not shown here to enhance readability

Fixed effects used for technological category of citing patent, country of citing patent, citing patent year and time lag between patents

** significant at 1%; * significant at 5% (In case of ratios, whether statistically different from 1 is tested)

Table 11: Cross-border learning effects from foreign innovative activity by MNCs

The regressions test if the probability of cross-border knowledge flows increases with increased presence of citing assignee in the cited country and/or with increased presence of cited assignee in the citing country. The sample used here consists only of the observations where the citing patent and cited patent originate in different countries, and where the citing entity is a home base.

Presence of citing assignee in cited country	0.030**
	(0.004)
	[0.16]
Presence of cited assignee in citing country	0.011**
	(0.004)
	[0.06]
Scale of citing assignee	-0.012*
	(0.006)
	[-0.06]
Scale of cited assignee	0.031**
	(0.005)
	[0.17]
Observations	3,027,928

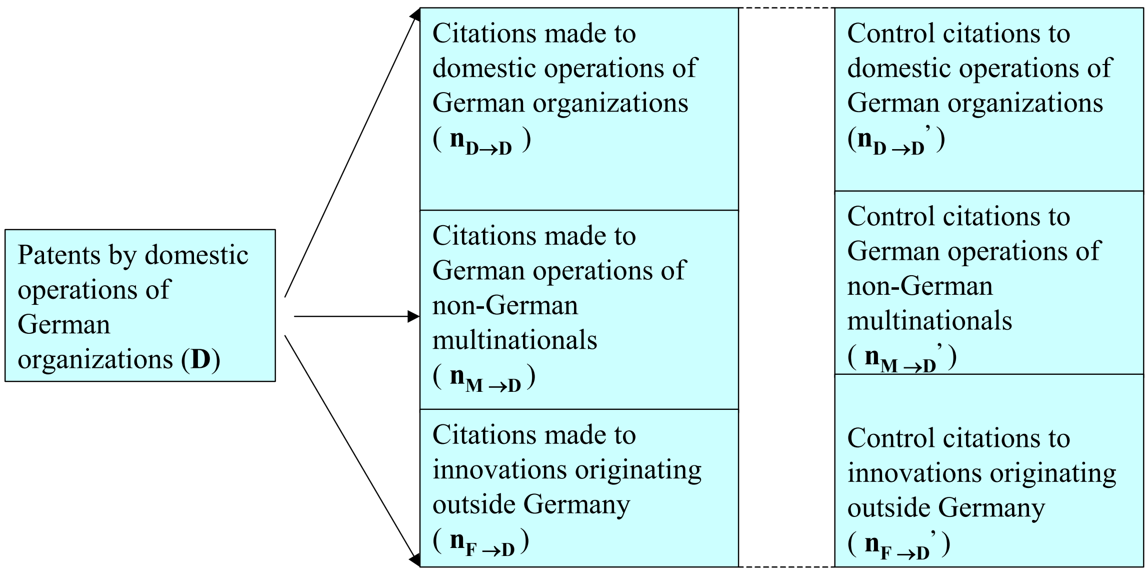
Robust standard errors in parentheses, with clustering on citing patent
Marginal effects in square brackets after multiplication with 1,000,000
Controls for technological similarity of citing and cited patent included in regression, but not shown here
Fixed effects used for technological category of citing patent, country of citing patent, citing patent year and time lag
** significant at 1%; * significant at 5%

Table 12: Mean citations to USPTO and EPO patents in USPTO data

This table lists the mean number of citations per patent made to European Patent Office (EPO) patents and USPTO patents by a sample of 1,612 USPTO patents from 1995. Identification of whether a cited EPO patent had an "equivalent" US patent was done using the "OECD Triadic Patent Families" database, which has information of patents filed for the same innovation at both USPTO and EPO. The means are reported for all patents, and then separately for patents arising from inventors based in the US and outside the US.

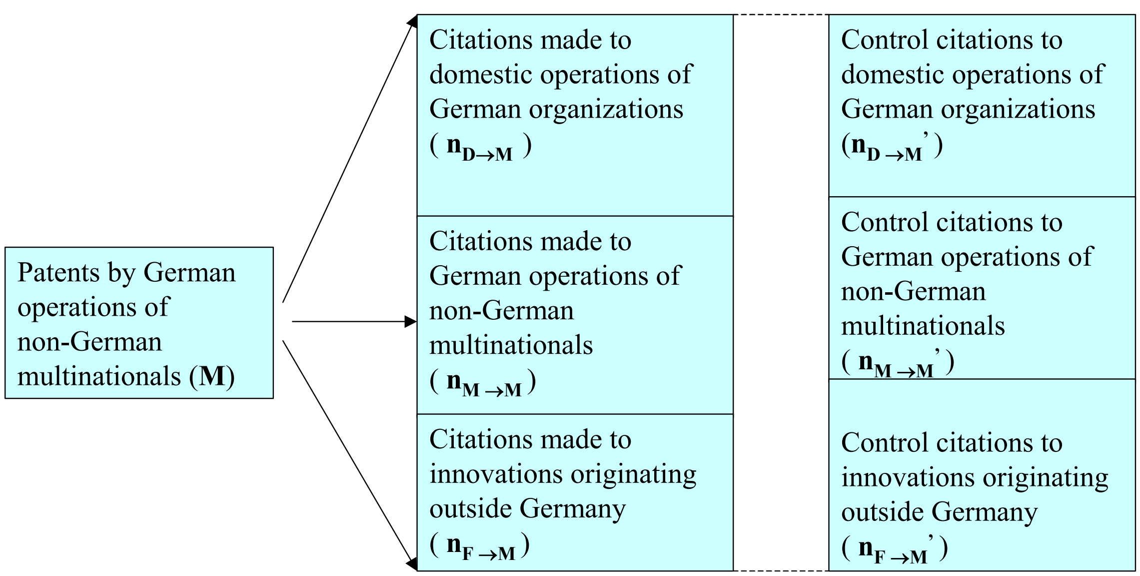
	Citing patents from all countries			Citing patents from US		Citing patents not from US	
	All assignees (N=1,612)	Domestic (N=810)	MNC (N=802)	Domestic (N=436)	MNC (N=369)	Domestic (N=374)	MNC (N=433)
Mean number of citations to USPTO patents	5.84	5.68	6.00	6.75	6.95	4.42	5.19
Mean number of citations to EPO patents	1.12	0.83	1.41	0.77	1.42	0.89	1.41
Mean number of citations to EPO patents with "equivalent" US patents in the OECD triadic database	0.32	0.22	0.43	0.24	0.39	0.21	0.46

Figure 1. Matching approach for detecting knowledge diffusion (using Germany as an example)



Comparison of $\mathbf{n}_{D\rightarrow D}$ with $\mathbf{n}_{D\rightarrow D}$' gives domestic-to-domestic (D→D) spillovers

Comparison of $\mathbf{n}_{M\rightarrow D}$ with $\mathbf{n}_{M\rightarrow D}$' gives multinational-to-domestic (M→D) spillovers



Comparison of $\mathbf{n}_{D\rightarrow M}$ with $\mathbf{n}_{D\rightarrow M}$' gives domestic-to-multinational (D→M) spillovers

Comparison of $\mathbf{n}_{M\rightarrow M}$ with $\mathbf{n}_{M\rightarrow M}$' gives multinational-to-multinational (M→M) spillovers