9-2005

Outward R&D and Knowledge Spillovers: Evidence Using Patent Citations

Ioana Popovici
Department of Economics, Florida International University

Outward R&D and Knowledge Spillovers: Evidence Using Patent Citations

Ioana Popovici
Florida International University

September 2005

FDI is believed to be a conduit of new technologies between countries. Many have studied the cross-border knowledge diffusion due to outward foreign direct investment (FDI), but this paper is the first to study the advantages of outward FDI for the home country of multinationals conducting R&D abroad. To address this issue, we use patent citations as a proxy for technology spillovers and we bring empirical evidence that supports the hypothesis that a U.S. subsidiary conducting R&D overseas facilitates the flow of knowledge between its host and home countries.

JEL Classifications: O30, O34, F23

Keywords: patents, spillovers, R&D, FDI

Department of Economics, Florida International University, Miami, FL 33199. email: Ioana.Popovici@fiu.edu. Helpful comments from Peter Thompson and Mihaela Pintea are gratefully acknowledged.

1. Introduction

A critical determinant of a country's economic performance is its ability to adopt new technologies. While relatively few countries have the capability to consistently develop new technologies themselves, all countries can potentially gain from identifying and exploiting technologies developed elsewhere. Unfortunately, empirical research strongly suggests that knowledge flows are geographically localized (Jaffe, Trajtenberg and Henderson, 1993, Thompson and Fox-Kean, 2005, Thompson, 2005, Agrawal, Cockburn and McHale, 2003, Agrawal, Kapur and McHale, 2004). That is why one source of competitive advantage for an economy is its favored access to local knowledge.

However, there is evidence that countries do have channels through which they can overcome the effects of geography. For example, Coe and Helpman (1994) show that the level of foreign R&D of its trade partners has a critical influence on total factor productivity, providing crude evidence that international trade can facilitate the transfer of knowledge. Foreign direct investment (FDI) may also facilitate the international flow of technology. Extending the theme that trade and factor flows are often substitutes, Keller and Yeaple (2004) demonstrate that economies also benefit from knowledge spillovers associated with inward FDI.

The existing literature largely focuses on knowledge flows from multinational corporations (MNCs) to their overseas subsidiaries and to other firms located in the host country. But MNCs also go overseas to acquire foreign knowledge and one might therefore expect that outward FDI facilitates knowledge flows to the home country of the MNC. One of the first indications that this type of technology flow occurs is provided by Globerman, Kokko, and Sjoholm (1996). Using data from patent citations, they examine the differences in technology sourcing by Swedish multinationals and non multinational firms. They find that outward FDI in a host country has a positive effect on the odds of that country being cited in Swedish patent applications of multinational and non multinational firms in Sweden. This could indicate that knowledge acquired by multinationals abroad diffuses to other firms from the home country of the multinational (in this case Sweden), as well.

This paper further investigates the extent to which domestic firms might benefit from MNCs outward investment. One area that seems particularly promising for generating knowledge flows arises when MNCs establish R&D-conducting subsidiaries abroad. These R&D laboratories overseas exploit local knowledge due to their involvement in local social networks. This knowledge is appropriated by the parent due to the formal ties between the MNC and its subsidiary. Domestic firms, in turn, access the overseas knowledge because they are part of the same network as the MNC. In summary, the establishment of an R&D affiliate overseas reduces the chain of social relations needed for knowledge to be passed from local firms in the host country to domestic firms in the home country of the MNC, creating a "small world" phenomenon.

This paper uses patent citation data to assess the evidence for such small world phenomena. It is a further step in understanding how knowledge diffuses internationally. But there are also policy implications. Cross-border knowledge flows create a positive externality for the multinational's country of origin, and the presence of such externalities must fundamentally affect our policy stance toward outward FDI investment, especially of the type that is designed to create knowledge through R&D conducted abroad.

Measuring knowledge flows through patent citations is fraught with danger. We have understood the need for an adequate control group against which to make comparisons since Jaffe, Trajtenberg and Henderson's (1993) seminal contribution. Thompson and Fox-Kean (2005) have shown how sensitive the results can be to the selection of the control group, and Thompson (2005) has provided an alternative that exploits differences in the degree to which citations added by examiners and citations added by inventors represent knowledge flows between inventors.

We subject our hypothesis that outward FDI in knowledge-creating subsidiaries stimulates knowledge flows from the host country to a particularly stringent test that combines the control-group methodology of Jaffe, Trajtenberg and Henderson (1993) with the examiner-inventor distinction proposed by Thompson (2005). We construct a dataset of patents assigned to U.S. companies that cite patents granted to U.S. affiliates abroad. For each of these patents we then compare its citations to patents granted to firms in the host country of the affiliate to its other citations. Specifically, we ask whether there is a difference in the probability that citations to the host country are more likely to have been added by the inventor (and hence more likely to represent a real knowledge flow) than are the other citations listed in the patent (most of which are to other U.S. firms). We then compare this probability difference with a control group, consisting of patents in the same technology classes that did not cite any U.S. affiliate in the host country. In short, we make a comparison between the citing and control groups of within-patent variations across groups of citations in the rate of examiner-added and inventor-added citations.

This seems like a rather stringent test for our data. Nonetheless, we find strong evidence that research-conducting subsidiaries stimulate knowledge flows to the home country. Citations to the host country are more likely to represent knowledge flows for U.S. firms citing the affiliate than for other U.S. firms that do not cite it. This finding is evident in the raw data, and it appears to be a robust feature of conditional logit regressions that control for a number of potentially confounding effects.

The layout of the remainder of the paper is as follows. Section 2 relates this work to prior literature. Section 3 describes the methodology and identification assumptions. Section 4 describes the data, Section 5 presents the results, and Section 6 offers some concluding comments.

2. Related Literature

As mentioned earlier, a central role in the theories of economic growth is played by technology advance and the way knowledge diffuses (Romer, 1990, Grossman and Helpman, 1991). A potential source of competitive advantage for an economy could be its access to new technologies generated in its geographic proximity. Alfred Marshall (1920) pointed out that the localization of knowledge spillovers is one of the three explanations for the geographic concentration of industries, along with specialized labor pooling and development of intermediate product industries. But measuring knowledge flows is a difficult task. Although Paul Krugman pointed out that “they leave no paper trail by which they can be measured and tracked”, the empirical research of Jaffe, Trajtenberg and Henderson (1993) indicate that “knowledge flows do sometimes leave a paper trail in the form of patent citations to prior art”. Their analysis tests the extent of spillover localization using patent citations and finds strong evidence of localized knowledge spillovers, even after controlling for agglomeration of industries. Using a different methodology, Thompson and Fox-Kean (2002) reassess the empirical work of Jaffe, Trajtenberg and Henderson (1993) and find evidence of international, but not intra-national localization of knowledge spillovers. Thompson’s (2004) estimation produces new evidence of international and intra-national localization of knowledge.

Other empirical studies strengthen Jaffe, Trajtenberg and Henderson (1993)’s finding that knowledge spillovers are localized. Agrawal, Cockburn and McHale (2003) develop a model based on social relations between inventors. They find empirical evidence that supports the hypothesis that inventors that move to a different geographic area continue to be a source of knowledge spillovers for their past locations’ innovative activity. Their result is relevant for determining what an economic location loses when skilled workers leave. Agrawal, Kapur and McHale (2004) examine the effects of co-ethnicity on knowledge flows, in a sample of Indian inventors. They find evidence that knowledge flows, besides being localized, are also ethnically sticky. They find that members of the Indian Diaspora are more likely to generate citations to other Indian inventors, even after controlling for the ethnic clustering of innovative activity. Another result of their paper is that co-location and co-ethnicity are substitutes, meaning that co-location becomes less important when there are alternative channels through which knowledge can diffuse.

In a world with international trade, international exchange of information, and foreign direct investment, besides the domestic R&D capital stock, a major contribution to a country's technological progress and productivity growth is brought by the foreign R&D capital stock. Coe and Helpman (1994) find empirical evidence that both domestic and foreign R&D efforts contribute to total factor productivity. Their empirical results suggest that FDI and international trade are important conduits of international knowledge diffusion. Keller and Yeaple (2004) find evidence that FDI in the U.S. accounts for more than 10% of U.S. productivity gains. The evidence they discover for imports-related knowledge spillovers is weaker. Head, Ries and Swenson (1994) study the location decision of a sample of Japanese manufacturing plants in the U.S. to examine the extent of agglomeration effects in industries. They find evidence that location choice is not driven solely by natural environment and specialized labor found in a particular state. Knowledge spillovers and intermediate inputs are at least as important explanations of agglomeration economies.

If knowledge is sticky and confined within spatial boundaries, firms' location is an important source of competitive advantage. But FDI can be crucial in overcoming this geographic constraint that firms face. MNCs can play a central role in the international diffusion of knowledge. There is a vast literature that deals with knowledge spillovers and MNCs. It was initially suggested that foreign direct R&D investment tends to be oriented towards supporting overseas manufacturing investments and customizing existing products for the special needs of local foreign markets. Mansfield, Teece, and Romeo (1979) find that U.S. R&D overseas is more focused on adjusting existing products to fit foreign markets than to developing new products and technologies. Recent studies, however, suggest a rapid growth of R&D investments overseas aimed at developing new product ideas. Florida (1996) uses regression analysis on R&D spending and a survey of foreign R&D labs in the U.S. to find that technology-oriented activities conducted by these laboratories are increasingly important, compared to market-oriented activities. His findings are confirmed by Kuemmerle (1999) in a detailed empirical survey of 32 large multinational corporations in 5 technological advanced countries.

Empirical studies emphasize knowledge flows from foreign MNC subsidiaries to host country domestic firms (Singh, 2004, Feinberg and Majumdar, 2001). But outward FDI, as well, can be a way to get access to foreign technology. It is one of the reasons why MNCs establish subsidiaries overseas, in industry clusters with leading technologies. Empirical studies suggest that there are significant knowledge flows from the host country domestic firms to foreign MNC subsidiaries (Almeida, 1996, Frost, 2001, Singh, 2004). Using data from the semiconductor industry, Almeida (1996) tests the hypothesis that knowledge flows from domestic firms in the U.S. are higher to foreign subsidiaries than to other domestic firms. Also, MNCs can facilitate knowledge diffusion through their internal structure and social networks. Singh, 2004 suggests cross border knowledge flows are higher within an MNC than between different firms.

Moreover, empirical studies try to explore what are the conditions under which foreign subsidiaries of MNCs are likely to use knowledge sources located in the home country of the firm or the subsidiary's host country to a higher extent, during the process of innovation. Frost (2001) argues that innovations in technological fields in which the host country has a technological advantage will be more likely to use technical ideas originating in the host country. Also, innovations that aim at customizing the existing products and technologies to local markets are more likely to draw upon technical ideas generated in the home country of the parent firm, rather than the host country of the subsidiary. Furthermore, the more autonomous the subsidiary is in the process of innovating, the less aligned it is to the existing knowledge of the home parent firm, the more likely it is to generate citations to patents originating in the host country. Moreover, subsidiaries that do not have the credibility and the resources to be accepted to participate in local knowledge-sharing networks depend more on the innovations generated in the home country. Since credibility takes time to be established, older subsidiaries tend to draw more upon technical ideas originating in the host country.

None of the earlier studies have focused on the flow of knowledge from the host country to the home country of the MNC's affiliates. In this paper, we analyze this type of spillovers.

3. Methodology

Although the analysis of patent citations has been the dominant way to address questions of knowledge flows, there are some questions related to the validity of drawing inferences about spillovers from them. One criticism is the extent to which citations represent knowledge flows, as not all citations represent spillovers and not all knowledge flows generate a citation. Second, empirical studies using citations have tried to control for the geographic concentration of industries by identifying control citations, but the selection of these comparison groups is challenging. This section describes our methodology and addresses how it tackles these questions.

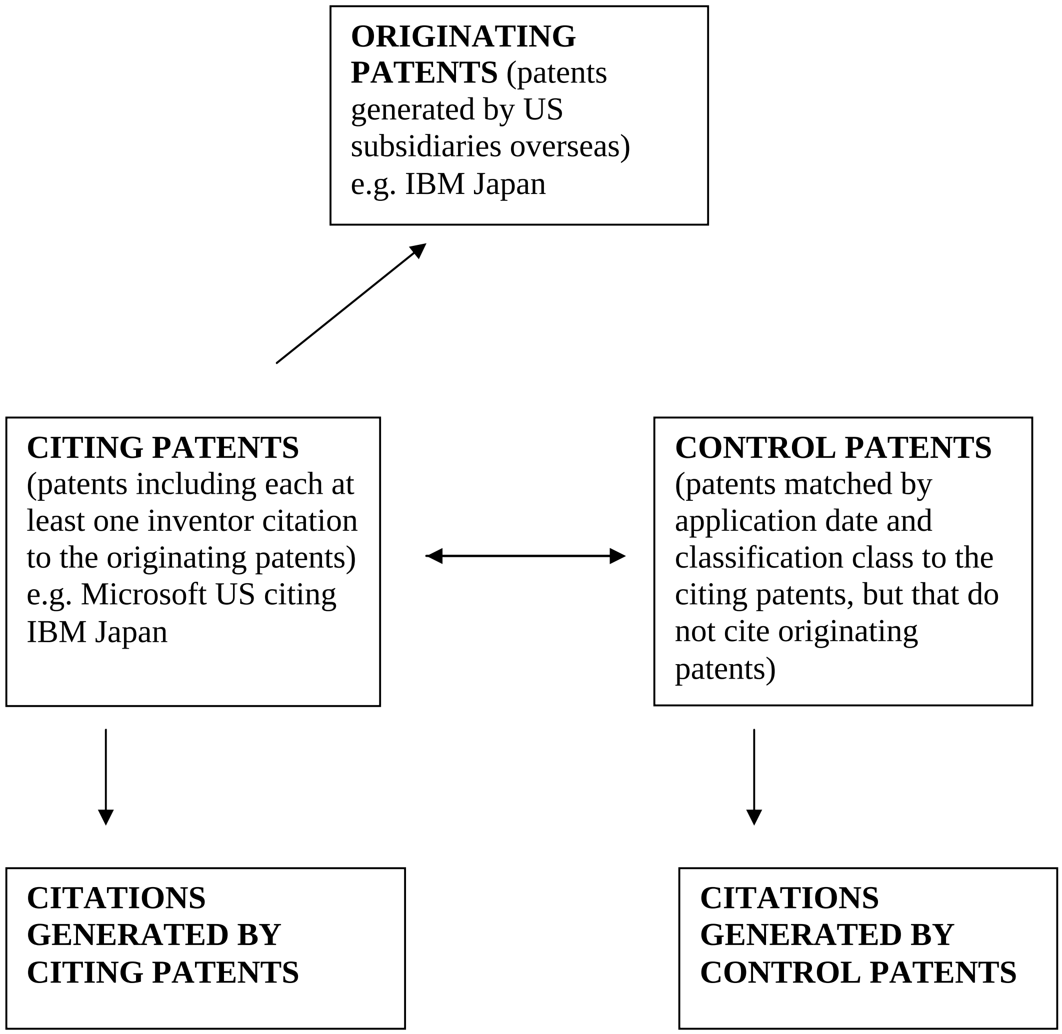
3.1. Patent citations to prior art as proxies for knowledge spillovers

The technological change literature has witnessed an increasing use of patent citations as a measure for knowledge spillovers. Jaffe, Trajtenberg and Fogarty (2000) or Duguet and MacGarvie (2003) are studies that assess the validity of interpreting citations as a proxy for technology diffusion. Jaffe, Trajtenberg and Fogarty (2000) surveyed inventors about their familiarity with the citations made by their patents, when and how they learnt about these citations. They find that only a quarter of the responses indicate a clear knowledge spillover, another quarter only the possibility of a knowledge flow, whereas half of the answers did not correspond to a spillover. Moreover, only about a third of the inventors had a high level of familiarity with their patents' citations. The authors suggest that this might be due to examiner included citations. Thus, the results of the survey suggest that patent citations can be used as a "noisy signal" of technology flow. Duguet and MacGarvie (2003) compare firms' citations with a survey that included questions about knowledge spillovers. They find support for the hypotheses that patent citations are highly, but not perfectly correlated with knowledge flows. According to Jaffe, Trajtenberg and Fogarty, 2000's survey, among the most important factors that add noise to patent citations are the examiner citations of patents that the inventors knew nothing about, thus not corresponding to spillovers, and the inventor citations of patents that do not indicate a knowledge flow (e.g. patents that the innovator learnt about after the completion of his invention and included them only to avoid litigation).

The most serious concern about using aggregate citations as a proxy for knowledge flows is the noise added by the citations generated by the examiners. Studies of the role of patent examiners in generating prior art (Cockburn, Kortum and Stern, 2002, Sampat, 2004, Thompson, 2004, Alcacer and Gittelman, 2004) suggest that the added noise is even higher than initially anticipated. Cockburn, Kortum and Stern (2002) and Sampat (2004) show that there is substantial heterogeneity among patent examiners which affects the patents' characteristics. Thompson (2004) finds strong evidence of localization of knowledge spillovers using an identification strategy based on the distinction between inventor and examiner added citations. He finds systematic differences among inventor and examiner citations geographic matching rates with the citing patents. Alcacer and Gittelman (2004) assess the validity of using aggregate citations as a proxy for knowledge spillovers. They compare inventor and examiner citations along a number of dimensions and find that examiner citations substantially bias the inferences made using aggregate patent citations. Although some inventor citations are not spillovers and a few examiner citations correspond to flows of knowledge (e.g. the innovator forgot to include them), overall, these studies suggest that inventor citations are more likely to indicate a knowledge flow than examiner citations. In other words, the probability that an inventor citation is a spillover is greater than the probability that an examiner citation represents a knowledge flow.

In this paper, we test the hypothesis that conducting R&D overseas by U.S. corporations is beneficial for other U.S. companies as well. They can take advantage now of the technology generated by domestic firms in the host countries of the subsidiaries. To test whether an affiliate overseas facilitates the diffusion of knowledge between the two countries, we analyze the likelihood of knowledge spillovers from the host country of a U.S. subsidiary to the U.S., conditional on a technology flow from the affiliate (e.g. the probability of a spillover from Japan to the U.S., conditional on the existence of a knowledge flow from a U.S. subsidiary in Japan). We construct a set of citing patents, each of them citing at least one originating patent of a U.S. affiliate overseas. Figure 1 illustrates the structure of the dataset used. To satisfy the condition that an actual flow of knowledge has occurred, we discard all examiner citations from this set. Based on the earlier observation that inventor citations are better signals of knowledge spillovers, we keep only the ones included by the inventor, as better signals of knowledge spillovers.

FIGURE 1. Dataset Structure



3.2. Controlling for other unobservable differences

We test whether there are significant knowledge flows from the host country to the U.S., conditional on the existence of technology flows from an affiliate. In other words, we test whether the presence of the subsidiary facilitates the flow of knowledge between its host and home country. We use the fact that an inventor citation is more likely to represent a technology spillover than an examiner citation. For U.S. patents that cite U.S. affiliates in foreign countries, we measure the likelihood that a citation to the host country is added by the inventor. In other words, we measure the probability that a host country citation is likely to represent a knowledge flow (e.g. for U.S. patents that cite U.S. affiliates in Japan, we measure the probability that citations to Japanese firms are added by the inventor). To make this measure meaningful, we have to compare it to a baseline. Countries differ greatly in their areas of technological specialization. Are knowledge flows between the host and home country of a subsidiary due to the presence of the affiliate in that country, or is it the mere interest of U.S. firms in a field in which the host country is specialized? We have to distinguish between the effects of the geographic distribution of technological activity and those due to the presence of the subsidiary. We have to control for the existing specialization of the host countries. If, after controlling for industry agglomeration, we still find significant knowledge spillovers between the host and home country of the affiliate, we are one step closer to finding evidence that these knowledge flows are, at least partly, due to the presence of the affiliate in that country.

Different methodologies have been used to control for the geographic agglomeration of activities by technological field. In studying the localization extent of innovative activity, Jaffe, Trajtenberg and Henderson (1993) compared the geographic matching rates of citing patents and their citations with the matching rates of cited and “control” patents. The latter were constructed to resemble the citing patents in technological class and application date, but to contain no citation to the cited patents. These controls were designed to serve as a benchmark, to make a distinction between the geographic localization of spillovers and the agglomeration of industries. Thompson (2005) developed a new method to control for the existing geography of production. He made use of a 2001 change in the reporting of patent data that shows whether the inventor or the examiner is responsible for including the citation.

To control for the technological specialization of countries, we follow the procedure developed by Jaffe, Trajtenberg and Henderson (1993). For each patent that contains a citation to a U.S. affiliate overseas, we find a control patent that does not cite a subsidiary in that host country, but has the same technological class and a close application date to the citing patent. We use this new set of control patents and compare our estimate for the set of citing patents with the same estimate for the control set. We compare the probability that a citation to the host country of the affiliate (e.g. a citation to a Japanese firm) is included by the inventor for patents citing the subsidiary (e.g. citing a U.S. affiliate in Japan) with the same probability for controls. In other words, we calculate the difference between the frequency of an inventor citation to the host country, conditional on a citation to a subsidiary located in that country, and the same frequency, conditional on the non-existence of such a citation. But what does this difference tell us? To be able to draw inferences, we have to compare this difference for the citations to the host country (e.g. citations to Japanese firms) to a benchmark that tells us what is the expected difference for the rest of citations in the dataset (those that do not refer to host countries firms, e.g. that do not cite Japanese firms). In other words, we try to eliminate other unobservable differences that could influence our inferences. If, even after controlling for other unobservable differences, we still find evidence of significant knowledge spillovers between the U.S. and the affiliates' host countries, we can conclude that the presence of the subsidiary is one of the factors that contribute to the international diffusion of technology.

An issue related to the examiner-inventor distinction is the observation that both types of citations may represent knowledge flows. An example is the case when the examiner adds a citation that the inventor failed to include, although he had knowledge about it. This sort of noise just makes it more difficult to find significant differences in citation patterns and our regressions will provide an underestimate of the true effects. However, as long as the probability that an inventor citation is a knowledge flow is greater than the probability that an examiner citation represents a knowledge flow, a sufficiently large dataset will allow us to identify the statistical significance, if not the true magnitude, of subsidiary-induced knowledge spillovers

3.3. Empirical specification

Estimation is carried out using Chamberlain's (1980) conditional logit regression. Therefore, we cannot obtain consistent estimates for the fixed effects. Also, the effects of independent variables that lack within group variation are not estimated. An alternative would be unconditional logit estimation. Although Katz (2001) has shown that the unconditional estimator has a negligible amount of bias when the number of observations per group, T, exceeds 16, in our sample T varies greatly. Although it averages about 48 per citing patent and 18 per control patent, there are many patents for which the number of citations is much smaller. There has been little research into the properties of the unconditional estimator in this case. According to Katz (2001), the conditional logit estimator is always unbiased regardless of the number of observations per group.

Also, although a fixed effects linear probability model would provide unbiased estimates, using OLS, heteroskedasticity will lead to unusable standard errors. Either constructing heteroskedastic-robust standard errors or using weighted least squares will suffer from the incidental parameters problem, as our sample has a large number of groups and a variable number of observations per group.

4. Dataset Construction

Table 1 shows summary information about the dataset. We began with a sample of 675 originating patents, consisting of all patents granted in 1998 to overseas subsidiaries of Fortune 50 companies. Unfortunately, the USPTO does not follow a rule as to whether to list the name of the subsidiary or the parent firm under the patent assignee. Therefore, to construct this originating set, we used the NBER Patent-Citations data file, that matches each patent assignee name to the name of the corporation as it appears in Compustat, hence to the name of the parent firm. We replaced the assignee name in the set with the Compustat name of the parent company. Then, following standard practice, the first inventor's country was taken to be the country where the innovation took place. Hence, patents generated by U.S. subsidiaries overseas are those assigned to U.S. companies and a foreign country of innovation.

We identified 3342 patents that cite one or more of the originating patents and were granted from January 2001 to present, hence, identifying whether their citations were made by the inventor or by the examiner. We kept only 2021 patents assigned to U.S. companies and inventors. Moreover, we discarded all citing patents whose citation to the originating patents was added by the examiner, since studies suggest inventor citations are a more accurate proxy for knowledge spillovers. This set of citing patents contains 894 patents generated by U.S. companies that contain an inventor citation for one or more of the 675 originating patents of U.S. subsidiaries overseas.

TABLE 1. Summary Statistics

	ORIGINATING PATENTS					
Number of observations	675					
	CITING PATENTS			CONTROL PATENTS		
Number of observations	894			894		
Citations per patent	48			18		
	CITATIONS BY CITING PATENTS			CITATIONS BY CONTROL PATENTS		
Number of Observations	42,981			15,306		
Fraction Inventor Added	0.908			0.713		
Fraction Subsidiary Host Citations	0.031			0.033		
	All citations	Inventor Citations	Examiner Citations	All citations	Inventor Citations	Examiner Citations
Mean File Date	Sep, 92	June, 92	Mar, 95	Aug, 92	Oct, 91	Nov, 94
Mean Issue Date	Oct, 94	July, 94	July, 97	Sep, 94	Oct, 93	Jan, 97

To construct the set of control patents, which will serve as a reference base, we identified a control patent for each patent in the set of citing patents. We followed a procedure developed by Jaffe, Trajtenberg, and Henderson (1993) and modified by Thompson and Fox-Kean (2004). For each citation of a U.S. company to a U.S. subsidiary overseas in the originating set, we identified a patent assigned to a U.S. organization, generated by U.S. inventors, with the same technology class and a close application date. We made sure the matched patent does not cite the originating patent or another patent generated by a U.S. subsidiary in the host country. The set of controls consists of 894 patents, one for each citing patent.

Then, for each citing and control patent, we used the front page bibliographic data for patents published by the USPTO to extract the following details for each citation made by these: assignee name and address, application and issue date, U.S. and international classification codes, primary examiner name, and name and address of inventors. After eliminating all citations granted before January 1, 1976, for which details are not available in machine-readable form, the set of citations generated by the citing patents consists of 42,994 observations. The set of citations made by the control patents consists of 15,307 observations.

There are some notable differences between the set of citing patents and the set of controls in terms of number of citations generated by each and fraction included by the inventor. The likely explanation is that this is due to differences in the average size of firms in the two groups. The citing patents draw upon technical innovations generated by subsidiaries abroad. Their assignees are generally large companies that tend to use a large number of past innovations, therefore generating more citations per patent. To account for this and other unobservable differences between the citing and control sets, we analyze within-patent variations. We compare the differences between the citing and control patents' citations to the host country with the differences between the citing and control patents' rest of citations (citations that do not refer to domestic firms in the host countries of the subsidiaries).

Table 2 shows that 90.7% of the citing patents' citations to the affiliate host country (e.g. citations to Japanese firms) are added by the inventors, whereas only 68% of the controls' citations to the host country are included by inventors. In other words, citing patents' citations to the host country are 22.7% more likely to be added by the inventor than controls' citations. Raw fractions show that citations to the host country are more likely to represent knowledge spillovers for the set of citing patents than for the set of controls. As mentioned earlier, we want to control for other unobservable differences between the two sets of patents, as well. Thus, we compare this difference with the same difference for the rest of the citations in the two sets.

The raw data show that for the subsidiary-citing patents there is no difference between the fraction of citations to the host country that are added by the inventor and the fraction of citations elsewhere that are cited by the inventor. That is, the likelihood that a citation represents a knowledge flow is the same regardless of whether the citation is to the host country or not. As most of the citations not to the host country are to patents granted to U.S. inventors, the raw data suggest that, if a firm cites a U.S. subsidiary abroad, knowledge in the host country is as accessible as knowledge at home. Contrast this with the raw data for the control patents. In this case, citations to the host country are less likely to be added by the inventors than are citations elsewhere. Although the difference is not statistically significant (the *z score* is 1.62), it provides a preliminary indication that knowledge flows from host countries are weaker when subsidiaries are not cited. Of course, the raw statistics may well be confounded by composition effects, and so we turn next to conditional logit estimations.

TABLE 2. Comparison of Inventor Citation Rates for Citing Patents and Controls

	Citing Patents Citations		Controls Citations	
Observations	42,994		15,307	
	Citing in affiliate's host	Not citing in affiliate's host	Citing in affiliate's host	Not citing in affiliate's host
Observations	1,320	41,661	513	14,793
Inventor Citations	1,198	37,840	349	10,569
Probability of an inventor citation	0.907	0.908	0.680	0.714

5. Results

The dependent variable, INVENTOR, is a dummy variable equal to one if the citation was added by the inventor and zero otherwise. The model contains four independent variables. The first is AFFILIATE, equal to one for all patents that cite a U.S. subsidiary abroad, and zero for control patents. The second is HOST, a dummy variable set equal to one for citations made to companies in the host country, and zero otherwise. For example, if patent X cites originating patent Y, a patent generated by a U.S. subsidiary in Japan, HOST = 1 for every citation that patent X makes to a Japanese firm and HOST = 0 for all other citations. The third variable is an interaction term between HOST and AFFILIATE, named INTERACTION in our model.

The coefficients on HOST and AFFILIATE provide the main results. Because international borders inhibit knowledge flows, we expect the estimated odds ratio for HOST to be less than one. The border is made more porous once a subsidiary is cited, so the odds ratio on INTERACTION is expected to be greater than one. The odds ratio on HOST alone therefore represents the extent to which international borders represent a barrier to firms that have not learned from a subsidiary abroad.[1] The product of the odds ratios for HOST and INTERACTION measures the impact of international borders on firms that have cited a subsidiary abroad. The raw data suggests that this product should not differ significantly from one.

We additionally expect that inventors' propensity to cite vary with time (Jaffe, Trajtenberg [1998], Thompson [2004]), so we include another explanatory variable, CITATION LAG, which measures the difference in years between the application dates of the cited and the citing patent. Finally, since we expect the inventors'

[1] Strictly speaking it provides something of an underestimate of this effect, because citations to the host country are compared with all other citations, and only a majority (rather than all) of these are citations to patents granted to U.S. firms.

propensity to cite in the same industry to behave differently from the propensity to cite across industries, we included another regressor, SAME CLASS, a dummy variable equal to one if the citing and the cited patent have the same technological first U.S. class, and zero otherwise.

We first run two regressions, a fixed effects logit and a logit with the same regressors, but without the fixed effects. Because the characteristics of citations are not necessarily independent of the citing patents, we use a citing patent fixed effects model. A Hausman specification test comparing the two specifications shows that, despite having selected control patents, there are significant citing patent effects, confirming the need to estimate conditional logit models.

Table 3 reports the results of the unconditional and conditional logit regressions. Coefficients for all variables are expressed as odds ratios. Column (1) summarizes the results of the unconditional logit. Columns (2)-(5) report the results of our attempt to address the heterogeneity problem by using a conditional logit with fixed effects per citing patent.

TABLE 3. Logit Estimates

(Odds Ratios Reported)

	(1)	(2)	(3)	(4)	(5)
	Unconditional Logit	Conditional Logit, Grouping on Citing Patent			
AFFILIATE	3.958 (55.27)	-	-	-	-
HOST	0.850 (-1.68)	0.776 (-1.74)	0.753 (-1.89)	0.782 (-1.67)	0.761 (-1.80)
INTERACTION	1.166 (1.12)	1.522 (2.28)	1.555 (2.34)	1.510 (2.22)	1.534 (2.25)
CITATION LAG			1.101 (23.67)		1.099 (23.13)
SAME CLASS				0.566 (-15.21)	0.586 (-14.11)
Number of observations	58,287	47,430	47,430	47,430	47,430

Z-scores in parentheses. CITATION LAG is measured as a deviation from the sample mean lag.

When estimating the conditional logit, 10,857 observations are lost due to a lack of within-group variation in the dependent variable. Column (2) shows results of the logit with two regressors. The variable of interest is the INTERACTION between HOST and AFFILIATE. Regressions (3) and (5) additionally control for the CITATION LAG, while regressions (4) and (5) additionally include SAME CLASS.

The results of the fixed effects estimations are almost identical across specifications. The (exponentiated) coefficient on HOST is less than one, significantly so at the ten percent level. As expected, citations to a set of patents that are all granted to foreign inventors are significantly more likely to have been added by the examiner. The coefficient of INTERACTION ranges between 1.51 and 1.55 and it is statistically significant at the five percent level. For patents citing subsidiaries, citations to the host country are 51 to 55 percent more likely to be inventor citations than are citations to the host country made by the controls, even after comparing to the rest of the citations in the dataset. The product of the odds ratios for HOST and INTERACTION is about 1.1 in each case, and this is not significantly different from one. Noting that these results are robust to the inclusion of controls for citation lags and technology class, the results seem like strong evidence for our hypothesis. Because inventor citations are more likely to represent knowledge spillovers than examiner citations, we can conclude that the presence of the affiliate abroad facilitates the international diffusion of knowledge.

The coefficient on CITATION LAG shows that, the older the cited patent compared to the citing one, the more likely it is that the inventor added the citation. Put differently, a citation to a ten year old patent is twice as likely to be added by the inventor. This result is consistent with Alcacer and Gittelman's (2004) estimate for vintage effects, as well as the finding in Thompson (2005). They also found that recent citations are more likely to be added by the examiners than by the inventors.

Again, coefficients for SAME CLASS are consistent with those estimated by Alcacer and Gittelman (2004). They show that citations to a different technological class are more likely to be added by inventors. Alcacer and Gittelman (2004) give two possible explanations for this. The first is that examiners are narrowly specialized by technological field and every patent is assigned by the "supervisory patent examiner" to a specific examiner (Cockburn, Kortum and Stern, 2002). The second draws attention to the fact that patents are classified during their examination. At that time, examiners search for relevant prior art and, sometimes, develop new classification codes.

As mentioned, the logit estimation with fixed effects per citing patent leads to a loss of 10,857 observations due to lack of within group variation in the dependent variable. Because we still need to control for systematic differences across citations, we follow Thompson (2005) and reduce the loss of effective observations by employing a conditional logit model with fixed effects per primary examiner of the citing patent. The trade-off here is a gain in sample size against the loss from using a broader fixed effect that may leave some significant within-group heterogeneity unaccounted for. As examiners differ greatly in their field of specialization, we expect that citations are not independent of the primary examiner of the citing patent. There are 812 examiners in our sample, therefore fewer groups than before, with an average of about 72 observations per group. As expected, only 3685 observations are lost by using this approach. Table 4 summarizes the results of primary examiner fixed effects logit estimation. Besides the regressors we had before, AFFILIATE is included. The coefficient could not be estimated with fixed effects per citing patent, as there was no within group variation in this independent variable. Specifications (2) and (4) replicate the results of the previous table, but the coefficient on INTERACTION is reduced markedly by the inclusion of CITATION LAG.

TABLE 4. Conditional Logit Estimates, Grouping on Primary Examiner

	(2)	(3)	(4)	(5)
AFFILIATE	4.815 (36.60)	1.694 (12.44)	4.838 (36.58)	1.668 (12.04)
HOST	0.824 (-1.50)	0.912 (-0.69)	0.817 (-1.55)	0.914 (-0.67)
INTERACTION	1.526 (2.53)	1.138 (0.79)	1.516 (2.47)	1.131 (0.75)
CITATION LAG		1.028 (9.13)		1.026 (8.64)
SAME CLASS			0.595 (-16.40)	0.873 (-4.47)
Number of observations	54,586	54,586	54,586	54,586

Z-scores in parentheses.

To check the robustness of the previous results, we replace the linear trend for the citation lag with citation lag fixed effects. Table 5 reports conditional logit results. The citation lag fixed effects are not reported in the table. (1) and (2) are citing patent fixed effects conditional logit regressions. Regressions (3) and (4) replace the citing patent with primary examiner fixed effects. The coefficient on INTERACT is a little smaller compared to the one obtained using the linear trend for the citation lag. It is also only marginally significant at 10% confidence level.

TABLE 5. Conditional Logit Estimates with Citation Lag Dummies

	(1)	(2)	(3)	(4)
AFFILIATE	-	-	4.995	4.659
			(35.94)	(34.67)
HOST	0.764	0.775	0.809	0.810
	(-1.76)	(-1.66)	(-1.58)	(-1.53)
INTERACTION	1.370	1.349	1.257	1.300
	(1.64)	(1.54)	(1.32)	(1.52)
SAME CLASS		0.578	-	0.641
		(-14.12)		(-13.88)
Number of observations	47,430	47,430	54,586	54,586

Z-scores in parenthesis.

6. Conclusions

Studies suggest that the level of foreign R&D, along with the domestic R&D stock, contributes to increases in productivity. Although numerous studies demonstrate that economies benefit from the knowledge spillovers associated with FDI, they focus on the benefits for the host country or the corporation conducting the R&D. Our study is one of the first to analyze the advantages of outward R&D for the home country. We take a further step in understanding how knowledge diffuses internationally. We study the technology flows between the home country (the U.S.) of MNCs and the host countries of their overseas affiliates, by using patent citations to prior art.

The most serious concern about using aggregate citations as a proxy for knowledge flows is the noise added by the citations generated by the examiner of the patent. Based on the evidence suggested by the existing literature that an inventor citation is more likely to represent a knowledge spillover than an examiner citation, we looked at within patent variation, making the distinction between examiner and inventor added citations.

A second issue was controlling for the existing technological specialization of the host countries. To tackle this issue, we followed a methodology that selects, for each citing patent, a patent that does not cite the originating one, but has close temporal and technological characteristics to the citing patent. We used these control patents as a benchmark.

We brought empirical evidence that supports the hypothesis that an overseas subsidiary of a U.S. MNC contributes to the diffusion of technology from the host country to the U.S. (e.g. a U.S. affiliate in Japan facilitates knowledge flows from Japan to the U.S.). We analyzed whether a U.S. company that cites a U.S. affiliate abroad is more likely to cite other host country firms than other U.S. firms that do not cite U.S. subsidiaries (e.g. we test whether U.S. firms that cite U.S. affiliates in Japan are more likely to cite Japanese companies). We find that citations to the host country are 51 to 55% more likely to be knowledge flows for U.S. firms citing the affiliate than for the rest of the U.S. firms, even after comparing to the rest of the citations in the dataset. Our results are evidence to the hypothesis that conducting R&D overseas can be benefic for the U.S. economy by facilitating the international flow of new technology from the host countries.

Our results are relevant in assessing optimal policies towards conducting R&D abroad and in determining whether policymakers should decide to subsidize overseas R&D or limit their subsidies for R&D conducted at home.

References

Agrawal, Ajay, Iain M. Cockburn, and John McHale (2003): "Gone But Not Forgotten: Labor Flows, Knowledge Spillovers, and Enduring Social Capital." NBER working paper 9950.

Agrawal, Ajay, Devesh Kapur, and John McHale (2005): "Defying distance: Examining the role of the Indian Diaspora." University of Toronto: Working paper.

Alcacer, Juan, and Michelle Gittelman (2004): "How do I know what you know? Patent examiners and the generation of patent citations." Stern School of Business, working paper.

Almeida, Paul (1996): "Knowledge sourcing by foreign multinationals: patent citation analysis in the U.S. semiconductor industry." Strategic Management Journal, 17(7): 155-165.

Cockburn, Iain M., Samuel Kortum, and Scott Stern (2003): "Are all patent examiners equal? The impact of characteristics on patent statistics and litigation outcomes." NBER working paper 8980.

Coe, David T., and Elthanan Helpman (1994): "International R&D spillovers." European Economic Review, 39: 859-887.

Duguet, Emmanuel, and Megan MacGarvie (2005): "How well do patent citations measure flows of technology? Evidence from French innovation surveys." Economics of Innovation and New Technology, 14(5): 375-393.

Feinberg, Susan, and Sumit K. Majumdar (2001): "Technology spillovers from foreign direct investment in the pharmaceutical industry." Journal of International Business Studies, 32(3): 421-438.

Florida, Richard (1996): "The globalization of R&D: results from a survey of foreign-affiliated R&D laboratories in the USA." Research Policy 26(1): 85-103.

Frost, Tony (2001): “The geographical sources of foreign subsidiaries’ innovations” Strategic Management Journal, 22: 101-123.

Globerman, Steven, Ari Kokko, and Fredrik Sjoholm (1996): “Technology Sourcing in Swedish MNEs and SMEs: evidence from patent data.” Stockholm School of Economics, Working Paper 125.

Grossman, Gene, and Elthanan Helpman (1991): Innovation and Growth in the Global Economy. Cambridge, MA: MIT Press.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg (2001): “The NBER patent citations data file: lessons, insights and methodological tools.” NBER working paper 8498.

Head, Keith, John C. Ries, and Deborah Swenson (1994): “Agglomeration benefits and location choice: evidence from Japanese manufacturing investments in the United States” Journal of International Economics, 38: 223-247.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993): "Geographic knowledge spillovers as evidenced by patent citations." Quarterly Journal of Economics, 108(3):577-98.

Jaffe, Adam B., Manuel Trajtenberg, and Michael S. Fogarty (2002): “The meaning of patent citations: report on the NBER/Case Western Reserve survey of patentees.” In A.B. Jaffe and M. Trajtenberg, Patents, Citations, and Innovations: A Window on the Knowledge Economy. Cambridge: MIT Press, pp. 379-401.

Katz, Ethan (2001): “Bias in conditional and unconditional fixed effects logit.” Political Analysis, 9(4):379-384.

Keller, Wolfgang, and Stephen R. Yeaple (2004): “Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States.” NBER working paper 9504.

Kuemmerle, Walter (1999): “Direct investment in industrial research in the pharmaceutical and electronics industries - results from a survey of multinational firms.” Research Policy, 28: 179-193.

Marshall, Alfred (1920): Principles of Economics. London: Macmillan.

Mansfield, Edwin, David Teece, and Anthony Romeo (1979): “Overseas research and development by US-based firms” Economica, 46(182):187-196.

Romer, Paul M. (1990): "Endogenous technological change." Journal of Political Economy, 98(5, pt. 2):S71-S102

Sampat, Bhaven N. (2004): “Examining patent examination: an analysis on examiner and applicant generated prior art.“ Georgia Institute of Technology, working paper.

Singh, Jasjit (2004): “Multinational firms and knowledge diffusion: evidence using patent citation data.” INSEAD working paper.

Thompson, Peter (2005): "Patent citations and the geography of knowledge spillovers: evidence from inventor- and examiner-Added Citations." Review of Economics and Statistics, forthcoming.

Thompson, Peter, and Melanie Fox-Kean (2005): "Patent citations and the geography of knowledge spillovers: a reassessment." American Economic Review, 95(1):450-460.