Borders and Distance in Knowledge Spillovers: Dying over Time or Dying with Age? - Evidence from Patent Citations

Yao Li*

Department of Economics, University of Western Ontario†

November 2009

Abstract

This paper uses a gravity framework to investigate the effects of distance as well as subnational and national borders in knowledge spillovers. Drawing on the NBER Patent Citations Database, we examine patent citations data at metropolitan level within the U.S. and the 38 largest patent-cited countries outside the U.S. Three key findings are documented. First, we find strong subnational localization effects at the Metropolitan Statistical Area and state levels: more than 90% of intranational border effects stem from the metropolitan level rather than state. This is consistent with the artifact of geographic aggregation at the state level for trade flows as in Hillberry and Hummels (2008). Second, border and distance effects decrease with the age of cited patent, which implies that new knowledge faces the largest barriers to diffusion. However, over time, border and distance effects are interestingly increasing. Finally, we find that (assignee) self-citations and aggregation bias are two sources of overestimated aggregate border effects of knowledge spillovers. While self-citations are only 11% of total citations, they account for approximately 50% of MSA and national border effects. Decomposing the data along geographic, age or industrial dimensions contributes to the reduction of border effects.

Keywords: Knowledge Spillovers, Gravity, Border Effect, Distance, Patent Citations
JEL Classification: F10, F29, O30, O33, O34

*My thanks to Ronald Jones, Raymond Riezman, Kala Krishna, Peter Egger and the participants of the CESifo Area Conference on Global Economy (Munich, 2008) for insightful comments and suggestions. I am particularly indebted to John Whalley, Jim MacGee and Hiroyuki Kasahara for guidance and inspiration. I also thank Carlos Rosell, Igor Livshits, Lance Lochner, Audra Bowlus, Karen Kopecky, Martijn van Hasselt, Nathan Sussman, Kun Peng, Philippe Belley, the participants of the Canadian Economics Association Annual Meeting (Vancouver, 2008), the Midwest International Economics (Columbus, 2008) and Macro Student Workshop participants at the University of Western Ontario for helpful discussion. All errors are mine.

†Correspondence: Yao Li, Department of Economics, University of Western Ontario, London, ON N6A 5C2, Canada (e-mail: yli267@uwo.ca).

1 Introduction

The degree of localization of intranational knowledge spillovers remains contentious. Recently Thompson and Fox-Kean (2005) have argued that only national boundaries restrict knowledge flows and that there is no strong evidence to support significant subnational barriers to knowledge diffusion. While Henderson, Jaffe and Trajtenberg (2005) and others argue that even intranational knowledge spillovers are indeed localized.[1] Also, the sources of localization of knowledge spillovers are not clear. Is this "nearby" effect more affected by physical distance or the national and subnational borders? How do borders and distance, as barriers to diffusion, affect knowledge flows? Are there any time trends or other profiles of these effects? The answers to these questions have significant implications for public policy on knowledge dissemination. Currently the importance of geographic proximity has attracted a lot of attention in knowledge spillovers literature, but differentiating the contribution from distance and borders and further analyzing their changing patterns and sources have not been explicitly investigated.[2]

In order to better understand the frictions affecting knowledge diffusion, the present paper asks three questions. First, how localized is intranational and international knowledge diffusion? To what extent do national borders, subnational borders and distance affect diffusion? Second, how does the pattern of knowledge diffusion change over time and with age? In this paper, "age" refers to the "age" of knowledge flows, defined by the citation lag between the citing and cited patents.[3] Third, what are the sources of border effects in knowledge diffusion?

To answer the above questions, we use a gravity framework to conduct a quantitative analysis of the magnitude of and the changes in the border and distance effects. We also attempt to tackle the "border puzzle" in the context of knowledge spillovers by examining the sources of overestimated aggregate border effects. We follow the principle assumption in the literature using patent citations that citations trace out knowledge flows: the fact that patents invented in region i cite patents invented in region j is equivalent to the fact that knowledge flows from region j to region i.[4]

[1] For example, Peri (2005) finds that pooled citations are strongly localized at state level within one country. Thompson (2006) and Alcácer and Gittelman (2006) find that inventor citations and examiner citations are both localized.

[2] For example, Thompson and Fox-Kean (2005); Henderson, Jaffe and Trajtenberg (2005); Thompson (2006); Griffith et al (2007) do not investigate distance. Only quite a few studies investigate distance explicitly in knowledge flows (Peri, 2005; Alcácer and Gittelman, 2006), but they either use dummy variables for distance intervals or drop internal distance, i.e., the distance from one region to itself is set to 0. No richer distance data have been investigated.

[3] Citation lag = the grant year of citing patent - the grant year of cited patent. For example, if patent A cites patent B which is 20 years old (i.e., B was granted 20 years ago), this is a relatively "old" knowledge flow, and the age of this knowledge flow is 20; if patent A cites patent B which was granted 2 years ago, this is a relatively "new" knowledge flow, and its age is 2.

[4] It should be noted that this paper only addresses the "pure" knowledge flows embodied in patent citations and

The advantage of using citations as a measure of knowledge flows is that citations leave a paper trail of knowledge flows (Jaffe, Trajtenberg, and Henderson, 1993), so they can provide interesting information tracking the direction and intensity of knowledge flows (see Section 3.3 for more detail). By differentiating citations by age, we characterize age distribution of different types of knowledge diffusion. We then estimate the subnational and national border effects as well as the distance effect for knowledge flows at aggregate level and by different criteria (age, category, and year). Based on those estimates, we analyze the changing patterns (age profiles and time trends) of the border and distance effects. Finally, we propose two sources of border effects in knowledge diffusion.

We use the NBER Patent Citations Data set of cross-patent citations (consisting of more than 3 million patents and more than 16 million citations) to study the border and distance effects in intranational and international knowledge flows across 319 MSAs (Metropolitan Statistical Areas) in the U.S. and the 38 largest patent-cited nations outside the United States. These regions cover more than 93% patents and citations in the NBER database between 1980 and 1997. We employ the metropolitan level data because the study of the geography of innovation shows that the majority of innovations are located in major cities indicating that innovation is an urban activity (Audretsch and Feldman, 1999, 2004). This raises doubts about the validity of large state border effect in previous literature.[5] The finer data set at the metropolitan level allows us to more fully explore the sources of subnational border effects and the nature of knowledge flow frictions.

Our findings support the strong subnational localization effects at the metropolitan and state levels. We find that more than 90% of intranational border effects stem from the metropolitan level rather than state. We also find that border and distance effects decrease with the age of knowledge. This finding suggests that, compared to older knowledge, new knowledge flows face more frictions, which is consistent with the nature of knowledge diffusion. However, over time, border and distance effects are interestingly increasing. Furthermore, we propose two sources of overestimated aggregate border effects of knowledge spillovers. One is self-citations, and the other is aggregation bias. Of total citations, only 11% are self-citations, but they account for approximately 50% MSA and national border effects.[6] Also decomposing data contributes to the reduction of the aggregate border effects.

This paper contributes to the emerging literature that explores the nature of knowledge diffusion using patent citation data. Currently most studies of knowledge flows do not explicitly differentiate

all knowledge studied in this paper refers to that associated with patents and citations since the general concept of "knowledge" contains extensive content and is difficult to quantify.

[5]Peri (2005) estimated that knowledge flows will be diminished to 20% when crossing state or province borders within one country.

[6]Self-citations refer to those citing patents and cited patents belong to the same assignee.

borders and distance in knowledge localization. Hence the contributing components of localization (for example, distance and internal distance; national borders and, especially, subnational borders, etc.) have not been well studied. Specifically, little is known about how subnational and international border effects and the effect of distance change in knowledge diffusion along different dimensions (time and knowledge age). The novel findings of this paper are the age profiles for aggregate border and distance effects of knowledge spillovers. The age profiles for friction factors in knowledge diffusion have not been previously reported. Also, our findings concerning the time trend of border and distance effects have not been extensively studied in the current literature.

This paper also contributes to the framework of the studies of knowledge spillovers, and in particular, subnational knowledge localization issues. The knowledge flow literature mostly exploits matching methodology, and it is difficult to reconcile the previous quantitative findings (e.g., Thompson and Fox-Kean, 2005; Henderson, Jaffe and Trajtenberg, 1993, 2005) due to the different criteria of selecting control groups.[7] Hence in this paper we use a gravity framework to avoid selecting control group and to estimate border and distance effects directly. Closely related work in empirical methodology is Peri (2005), who employs the gravity-like equation to study knowledge flows, using the subnational patent citation data at the state (or province) level. Peri's findings suggest large state border effects.[8] We use a finer, newly constructed data set at the MSA level to show that most subnational border effects exist at the metropolitan level, rather than at the state level. Knowledge diffusion is much more localized than we expect. When the MSA border is considered, the state border effect is very small.

Finally, we contribute to a large literature on gravity application and border effects. This paper presents the compelling empirical evidence for the resolutions of the border puzzle in knowledge flows. Part of the proposed resolutions might be extended and linked to border effects in trade flows. For example, when we decompose data from state level to MSA level, the state border effect is substantially reduced; if we further use disaggregated data at the category level, some state border effects are not significant at all. This is consistent with the findings of Hillberry and Hummels (2005), who argue that the state level home bias in trade flows is largely artifact of geographic aggregation.

The remainder of the paper is organized as follows. Section 2 briefly reviews the relevant literature on knowledge spillovers and border effect. Section 3 sets out the basic framework of analysis and details the empirical specification and data. Section 4 presents main results and

[7] Matching method was first used by Jaffe, Trajtenberg, and Henderson (1993) to study the geography of knowledge flows using patent citations. They matched each citing patent to a non-citing patent, which shares the same location with the citing patent, so as to control for the existing concentration of knowledge production.

[8] Peri (2005) estimates that only 20% of average knowledge is learned outside the average region of origin, i.e., there is around 80% of initial knowledge flows would be lost when they cross state border.

section 5 examines robustness. Section 6 concludes.

2 Literature Review

2.1 Knowledge Spillovers Literature

The last two decades have seen the development of a significant body of research on knowledge spillovers or knowledge flows. It is useful to distinguish between two branches of literature, one which focuses on measurement issues and another which focuses on the study of knowledge flows.

Measuring knowledge flows in a consistent, systematic way is a difficult task. Currently the first branch of literature contains three main measures of knowledge flows: R&D expenditures, royalties and license fees, and patent citations. Moreover, some alternative approaches consider trade flows or foreign direct investments as proxies for knowledge flows.[9] We briefly review the three main measures as follows.

First, some studies examine the spillover effect of international R&D on domestic productivity. The weakness of this approach is in distinguishing the effect of "pure" knowledge flows from the effect of technology flows embodied in advanced capital goods sold from one country to another (Jaffe and Trajtenberg, 1998). Another problem is finding an appropriate way to weight the foreign R&D. For example, Coe and Helpman (1995) find potent international R&D spillovers using trade volume weighted foreign R&D. But Keller (1998) challenges their results using the Coe and Helpman (1995) database, by weighting foreign R&D with randomly created trade patterns. At the firm level, Branstetter (1996) finds a strong intranational spillover effect but very small or even negative international knowledge spillovers, using technology proximity to weight other firms' R&D. This implies that R&D expenditures as a proxy for knowledge flows usually diffuse only within an economy and not across national borders. We need a direct and explicit measure of knowledge flows other than R&D expenditures if we want to investigate international and intranational knowledge spillovers simultaneously.

Second, using the international payments and receipts of royalties and license fees provides a precise measure of the value of knowledge flows. But so far there are no bilateral data with wide international coverage. The intranational data is even more difficult to find than international data. Only aggregate data for a few countries or firm-level data within a very restricted scope are available. For instance, Giummo (2003) examines the royalties received by the inventors/patentholders at nine major German corporations.

[9] See Peri (2005) for a brief review of this literature.

Third, using patent citations can give a direct paper trail of knowledge spillovers across different types of boundaries. Griliches (1990) and some other seminal works (for example, Jaffe, Trajtenberg and Henderson, 1993) started this line. Aggregated citation flows have recently been extensively used as proxies for knowledge spillover intensities. At the firm level, Hall, Jaffe and Trajtenberg (2005) explore the usefulness of patent citations as a measure of the "importance" of a firm's patents, as indicated by the stock market valuation of the firm's intangible stock of knowledge.

The second branch is the literature on the nature of knowledge flows using patent citations. Most studies focus on the geographic or institutional determinants. However, when they investigated geographic determinants, usually only geographic units (proximity) were examined without explicit distance measures. For example, Jaffe, Trajtenberge and Henderson (1993) and Jaffe and Trajtenberg (1998) find that citations are geographically localized. Inventors in the same country are 30 to 80% more likely to cite each other than foreign inventors. Griffith et al (2007) examine the home bias of international knowledge spillovers as measured by the speed of patent citations between countries and find that home bias is stronger in the pre-1990 period than the post-1990 period. Similar to previous studies using matching methodology, Griffith et al (2007) employ econometric duration models with fixed effects, in which a distance measure is not exploited. Recently only a few papers investigate distance and borders at the same time, but they do not capture the changing patterns and potential sources for these effects. For instance, Peri (2005) estimates the percentage of knowledge learned outside the region of origin using the data from subnational (state and province) regions in Europe, Canada and the United States. Compared to international spillovers, intranational knowledge flows are relatively less studied and the localization of intranational knowledge flows remains contentious in the literature.

2.2 Border Effect Literature

Another root of relevant literature stems from the large border effect in international trade, which remains a key puzzle in this field. Obstfeld and Rogoff (2000) refer to the "McCallum Home Bias in Trade" puzzle as one of the six leading puzzles in modern international macroeconomics. Anderson and Van Wincoop (2003) develop a theoretical gravity model to correct the bias in McCallum's (1995) estimates. In this paper, building on the gravity framework by Anderson and Van Wincoop (2003), we derive a gravity equation of knowledge flows (see Appendix) to investigate border and distance effects in knowledge diffusion. We use fixed effects estimation method as in Anderson and Van Wincoop (2003) and Feenstra (2002).

3 Empirical Specification and Data

3.1 Basic Framework of Analysis

We employ a gravity framework to analyze the friction factors in knowledge flows because we want to investigate the contribution of different types of borders and distance in knowledge diffusion, rather than the combined "localization effect". As stated in Combes (2008), "in addition to the reliable estimates of the impact of distance they lead to, the success of gravity models is due to their great explanatory power for flows, and this holds true whatever the geographical scale (countries, large or small regions), the period of study or the goods considered".

Let c_{ij} denote how many citations region j receives from region i, i.e., the number of citations region i makes to use the existing knowledge created by region j's patents, or, the quantity of knowledge flows from j to i. Let y_j be the total number of citations region j receives from all regions in the world. Hence y_j captures the size of region j's knowledge production capacity. "Region" is defined flexibly in this paper, using MSAs within the U.S. and 38 countries outside the U.S. We use subnational borders, national borders, distance and internal distance to proxy for the friction factor t_{ij} in knowledge flows between region i and region j. We follow the convention in gravity literature in hypothesizing that t_{ij} is a loglinear function of observables, bilateral distance d_{ij}, and whether there is a national border B^n_{ij} (1 if crossing countries, 0 otherwise), a state border B^s_{ij} (1 if crossing states within the U.S., 0 otherwise) and a MSA border B^m_{ij} (1 if crossing MSAs within the U.S., 0 otherwise). Other factors can also be added to knowledge flow frictions, such as adjacency and linguistic identity. Here we have chosen borders and distance for simplicity as well as to stay as close as possible to Anderson and Van Wincoop (2003), so that potentially we are able to compare the frictions in trade flows and knowledge flows under a common framework. Building on Anderson and Van Wincoop (2003), we derive a theoretical gravity equation of knowledge flows: (see Appendix for the derivation)

$$\ln(\frac{c_{ij}}{y_i y_j}) = k + \alpha ln d_{ij} + \beta_1 B^m_{ij} + \beta_2 B^s_{ij} + \beta_3 B^n_{ij} + \ln(Q_i)^{\sigma-1} + \ln(Q_j)^{\sigma-1} + (1-\sigma)\varepsilon_{ij} \tag{1}$$

where σ is the elasticity of substitution between all knowledge products (patent citations); k is a constant; Q_i and Q_j are quantity indices, referring to the measures of "multilateral knowledge flow resistance" variables as they depend on all bilateral resistances t_{ij}. Equation (1) is the theoretical gravity equation where we start our empirical work.

3.2 Empirical Gravity Equation

Equation (1) is hard to estimate since the multilateral resistance terms are not observable. We have two ways to handle this problem. One is to use region-fixed effects terms in place of the region-specific multilateral resistance terms; the other is applying structure estimation by constrained nonlinear least squares, as in Anderson and Van Wincoop (2003). In the current paper we use the fixed-effects estimator for the following reasons: (1)reducing the computation intensity; (2)leading to consistent estimates of model parameters (Hummels, 1999); (3)giving similar results to structure estimates (Anderson and Van Wincoop, 2003); (4) the fixed-effects method produces consistent estimates of the average border effect (Feenstra, 2002).

We use region-specific terms to control for the unobserved multilateral resistance terms. The empirical gravity equation then becomes

$$\ln(\frac{c_{ij}}{y_i y_j}) = k + \alpha ln d_{ij} + \beta_1 B_{ij}^m + \beta_2 B_{ij}^s + \beta_3 B_{ij}^n + r_1^i CI^i + r_2^j CE^j + (1-\sigma)\varepsilon_{ij} \quad (2)$$

where CI^i is equal to 1 if i is the citing region (destination region of knowledge flows) and 0 otherwise, and CE^j is equal to 1 if j is the cited region (source region of knowledge flows) and 0 otherwise. In general, the fixed effects control for any citing- and cited-region-specific characteristics. This is our baseline regression for cross-sectional data. We also construct the panel data to identify the time trend of effects of borders and distance.

3.3 Description of the Data

Patent and citation data originate from NBER Patent and Citation Database, which is publicly available and described in detail by Hall, Jaffe and Trajtenberg (2001). This data set contains all the patents (more than 3 million) granted by the U.S. patent office (USPTO) and, since 1975, all citations (more than 16 million) made by each patent of other patents, in which more than 40% patents granted to foreigners and more than 40% citations generated by foreigners (see Figure 1).

The most useful information is the inventors' geographic location by their registered residence and citations made and received by each patent. In the data set, we can identify in which country the inventor is located. If the inventor resides in the U.S., we also know in which state the inventor resides. Furthermore, we want to locate each patent at the MSA level. Among all inventors, 15% of them report the zip code of their residence in the U.S. and all inventors report the town/city or place name of their residence. We first locate inventors to MSAs by zip code and then locate the rest by town/city or place name. The matching was done using correlation files provided by the Office of Social and Economic Data Analysis (OSEDA) of the University of Missouri. We use MSAs as defined by the U.S. Census Bureau in 1990.[10] We also created 49 phantom MSAs, one for each state (except for New Jersey), containing all locations in non-metro areas.[11] Finally, we matched more than 93% U.S. inventors to 319 MSAs.[12] Then the region of a patent is denoted by the residence of its first inventor.[13] For a patent invented within the U.S., the region is the MSA of its location. For a patent invented outside the U.S. (we call it a "foreign" patent), the region is the country of its location. If a region i's patent cites a region j's patent, we assume that there is "one" knowledge flow from j to i at the first glance. Then, we aggregate the quantity of bilateral citation flows between each region pair ij every year as a measure of knowledge flows.

We use patent citations to measure knowledge flows for several reasons. Patents embody new ideas associated with knowledge. A patent awards to inventors the right to exclude others from the unauthorized use of the disclosed invention. The applicant has the legal duty to disclose any knowledge of the "prior art" hence citations to previous patents are included in the patent documents. Intuitively speaking, if patent B cites patent A, it implies that patent A represents a piece of previously existing knowledge upon which patent B builds, and over which B cannot have a claim. When patents generate citations, they leave a paper trail of knowledge flows (Jaffe, Trajtenberg, and Henderson, 1993). Thus, *patent citations*, rather than the patent stock itself, can provide interesting information tracking the direction and intensity of knowledge flows. Previous studies also find that the estimated value of a patent is correlated with subsequent citations, and that the most highly cited patents are very valuable (Giummo, 2003). This further suggests that patent citations is a good measure of knowledge flows.

We choose the sample of citations between 1980 and 1997 associated with each citing and cited patent pair whose inventors are residents of one of the 357 regions (319 MSAs within the U.S. and other main 38 countries). We choose the other 38 countries by their rank of knowledge production as well as the importance of their economy.[14] The time of citation is defined by the grant year of the citing patent. The cited patents in the sample are restricted to patents granted after January 1, 1976. Our final sample covers more than 93% patents and citations between 1980 and 1997 in

[10]The definition of MSAs evolves over time and there is slight difference between the definition of MSAs in 1990 and in 2000. We choose the definition in 1990 since our sample period is 1980-1997.

[11]In our sample, no citations come from non-metro area of New Jersey.

[12]These 319 MSAs include 270 MSAs as defined by the U.S. Census Bureau in 1990 and 49 artificial MSAs.

[13]The rule of "location by the first inventor" is designed by the constructor of NBER Patent and Citation Database.

[14]The sample (except for U.S.) is constructed by the following procedure: First, rank all countries by the total number of citations production and the total number of patents production, and choose the 30 largest countries in both ranking list. Second, use the intersection of these two groups of 30 largest countries. Third, plus all other OECD countries (which are not included in the first set except for Slovakia). Fourth, plus the OECD Non-Member Economies (China, Russia, Brazil) and India.

the world, which contains more than 1.6 million patents and more than 6.6 million citations. The present sample is more comprehensive than other recent knowledge flow studies.[15]

Table 1 presents the largest 20 countries by knowledge production capacity in terms of number of citations received. Not surprisingly, the U.S. ranks as the most productive and innovative country in the world. Japan, Germany, the United Kingdom and France are also at the top of the table. Table 2 reports some characteristics of the most and least innovative regions in our sample. The largest innovator is Japan, which receive more than 59,000 citations per year. We find that usually most innovative MSAs are those crossing multiple states. Hence it is useful to investigate state border and MSA border separately. The bottom of the list is occupied by Turkey, Iceland, and some low-cited MSAs, each with very small number of received citations. Usually the least innovative regions in the U.S. are located restrictively within one state.

Distance data come from CEPII's worldwide geographical database for countries and we use geodesic distances which are calculated following the great circle formula, using latitudes and longitudes of the most important cities/agglomerations (in terms of population). Within the U.S., we use coordinates of the largest city (by 1990 population) to locate MSAs. We also use the area-based internal distance formula to investigate the intra-regional knowledge flows (Head and Mayer, 2002).[16]

4 Main Results

We present in this section the estimates from equation (2) with different specifications to solve the previous three questions. We find that subnational border effect mainly comes from MSA level, rather than state level. We also find that movement of these friction factors (border and distance effect) in knowledge flows is falling with the age of knowledge but rising over time. Furthermore, we propose the compelling empirical evidence for the resolutions of border puzzle in knowledge flows by examining the sources of overestimated aggregate border effects.

[15]For instance, Juan Alcácer and Michelle Gittelman (2006) uses the sample of 1,456 patents and 16,095 citations; Peri (2005) uses the sample of 1.5 million patents and 4.5 million citations; Griffith, Lee and Van Reenen (2007) uses approximately 2.1 million cited patents.

[16]An often used measure of average distance between producers and consumers in a country, see Head and Mayer (2002), *Illusory Border Effects*, CEPII Working Paper No. 2002-01. We follow their formula: $d_{ii} = 0.67(area/\pi)^{1/2}$ in the context of flexible "region" to calculate the internal distance. Hence in our sample, $d_{ii} \neq 0$.

4.1 Basic Estimates of Border and Distance Effects

Table 3 is the basic estimates of border and distance effects for the whole sample (357 regions and 18 years) on aggregate knowledge flows. All coefficients are significant at the 1% level. To interpret the economic meaning of those coefficients, we take specification (1) as an example. For the whole sample, the distance effects are approximately -13% in the period of 1980-1997, which means that the knowledge flows will decrease 13% associated with a 1% increase in the distance holding everything else constant. In other words, halving distance will increase knowledge flows by 6.5%. Distance effect in knowledge flows is much smaller than that in trade flows. Halving distance increases trade by approximately 45% (Disdier and Head, 2006). This implies that knowledge flows are less affected by physical distance than trade flows. To examine border effects, we need to use the exponential formula. Specification (1) shows that, the intranational knowledge flow is 13.32 ($= e^{2.589}$) times higher than cross-nation-border knowledge flow; the intra-MSA knowledge flow is 8.45 ($= e^{2.134}$) times higher than cross-MSA-border knowledge flow; and the intra-state knowledge flow is 1.25 ($= e^{0.224}$) times higher than cross-state-border knowledge flow. Here we use the average border effects which is calculated as the exponent of the (absolute value of the) coefficient on the border indicator (Feenstra, 2002).[17] In other words, national border effect implies that 92.5% (= 1- $e^{-2.589}$) of initial knowledge flow is lost passing the country border, holding all other factors constant; 88.2% (= 1- $e^{-2.134}$) knowledge flow is lost crossing the MSA border; 20.1% (= 1- $e^{-0.224}$) knowledge flow is lost crossing the state border. We can see that MSA and national border effects are very significant, and substantially impede knowledge flow. The magnitude of state border effect is very small compared to the other two borders. On average, national border effect is larger than MSA border effect, and MSA border effect is much larger than state border effect. However, for aggregate knowledge flows, state border effect is still statistically significant.

When we use different specifications with MSA border and year effect included (see specification (1) and (2)), we find that the coefficients for log distance, MSA border and national border are quite stable, which belong to [-0.13, -0.15], [-2.13, -2.25], and [-2.43, -2.59] respectively. Dropping state border does not change the results much. It implies that MSA border captures most of intranational border effect in knowledge spillovers. However, if we only use state border to represent intranational border effect as in specification (3), we find that the magnitude of border and distance effects is much different with that in previous specifications. This implies some artifact of geographic aggregation at state level(see Section 4.3 for details). Also including year dummies substantially improve the estimation results and it implies that year heterogeneity is significant in the panel.

[17]Feenstra (2002) proves that this simple method can produce the consistent estimates with the structural estimates in Anderson and Van Wincoop (2003).

Another issue here is whether we use the normalization of dependent variable. Following Anderson and Van Wincoop (2003) and other previous literature on estimating border effect, we prefer the normalization method. Normalization makes the regression more robust. With normalization, the distribution of dependent variable will shape better than without normalization (see Figure 2 and 3). We also estimate the border and distance effects without normalization, i.e., treating lny_i and lny_j as independent variables. We find that the magnitude of border and distance effects does not change much.

We also consider the impact of self-citations on border and distance effects because presumably self-citations represent transfers of knowledge that are mostly internalized (Hall, Jaffe and Trajtenberg, 2001) but they are not necessarily locked in the same location. Hence investigating self-citations has important implications for the study of barriers to knowledge spillovers. Table 4 presents aggregate border and distance effects with and without self-citations. We find that self-citations partly exaggerate border and distance effects. After excluding self-citations, 85% of initial aggregate knowledge flows will be lost crossing national borders; 78% will be lost crossing MSA borders; and 12% will be lost crossing state border (see specification (7)). Including self-citations approximately doubles the aggregate MSA and national border effects, but does not change the order of importance of three types of border effects.

We use specification (7) as our baseline regression since most previous studies exclude self-citations and it is more convenient to compare the estimates without self-citations to previous literature. Peri (2005) excludes self-citations and finds that only 20% of average knowledge is learned outside the state (or province) of origin, i.e., 80% of initial knowledge is lost crossing the state border. The magnitude of the state border effect in Peri (2005) is similar to our MSA border effect. We show that it is not a true magnitude of state border, and 92% (=0.78/0.85) of intranational border effects come from metropolitan level, rather than state. Peri (2005) also finds that national borders diminish knowledge flows to 9% of the initial level. Our estimates show the relatively smaller national border effect, and there are still 15% of initial knowledge which can spill over to other countries.

4.2 The Changing Patterns of Border and Distance Effects

One might think that new knowledge and old knowledge might be different in diffusion. Hence, we expect that the different types of knowledge flow (e.g., international, intra-state, and intra-MSA, etc.) have different age distribution. We draw on the proportion of citation received in its total (lifetime) citations at each age to characterize the age distribution for each type of knowledge flow (see Figure 4 and 5). Figure 4 shows that there is approximately a 5-year lag between local and non-local knowledge flows, within-MSA and cross-MSA flows, as well as intranational and international flows.[18] Figure 5 presents the age distribution of knowledge flows without self-citations. By comparing Figure 4 and 5, we find that excluding self-citation substantially reduces the gap between the age distribution lines of local and non-local, within-MSA and cross-MSA as well as intranational and international knowledge spillovers. It suggests part of the border effects from self-citations. This has been confirmed by our previous estimates.

Another message conveyed by Figure 4 and 5 is that border and distance effects are expected to decrease with the age of knowledge since the integrals of the different age distributions converge with the age of knowledge. To verify this prediction, we decompose the whole sample to 5 subsamples by age group, using 5 years as an interval. The results are very significant as we expected and are presented in Table 5: distance and border effects are decreasing with age of knowledge. Hence new knowledge flows face the largest distance and border effects. The only exception is the state border effect. For old knowledge groups (more than 15 years old), state border effects are not significant and slightly deviate from the decreasing age profiles. However, the age profiles for MSA and national borders as well as for distance effect are very significant (all at 1% level). Also, we find that on average, national borders effect is larger than subnational border effect, and this holds true for each age group.

The age profiles of border effects and distance effect is not a surprising result and it is consistent with the nature of knowledge diffusion process in the real world. But the current literature abuses the geographical localization effect and usually use "time" instead of "age". When people argue that over time the tacit information embodied in knowledge is codified and is more easily to be transmitted across distance or borders, they actually mean over the age of knowledge. If we seriously differentiate the time effect and age effect, our findings suggest that new patents have a larger number of local citations than older patents. This would seem to make sense – new patents may be cited more often initially either by their owner (since they be part of an ongoing research agenda) or they may be known to other local firms/researchers before their formal patenting - which again would give local researchers/firms a head start. We think that "age" might be a good dimension complementary to "time" in examining the changing pattern of barriers in knowledge diffusion since knowledge diffusion involves two parties - the predecessor (cited) and the successor (citing), and "age" can capture the impact of the lag between these two sides. Only looking at the changing pattern over time might omit important information in knowledge transmission.

[18] Local knowledge flows refer to all intra-region flows, i.e., intra-MSA flows within the U.S. and intranational flows within a country outside the U.S. Within MSA and cross-MSA flows are specific to the knowledge within the U.S.

Now we turn to time trends of border effects and distance effect. In recent trade literature, whether distance is dying over time is an interesting topic which has already attracted lots of attention. However, in knowledge flow literature, only quite recently have economists started to concern this question (e.g., Griffith et al, 2007), and some conjectures are proposed which need serious empirical work to verify. For example, Henderson, Jaffe, and Trajtenberg (2005) proposed that localization effects are likely to fade over time, but they didn't give empirical evidence to support this conjecture in that paper.[19] In this paper we can investigate this issue through border effects and distance effects. So far no researchers investigate the time trend of border effects, probably due to the lack of a common framework of analysis and the difficulty of accessing the relevant data.

Figure 6 and 7 show the time trends of border and distance effects based on cross-sectional estimates for each year in our sample period (1980-1997). MSA and national border effects as well as distance effect are all increasing over time, while state border effects are very small and almost flat. For some years, the state border is not statistically significant. However, all MSA and national borders as well as distance are significant for each year. Again, border and distance effects with self-citations are always larger than those without self-citations, and national border effects are larger than subnational for each sample.

Why are border and distance effects increasing over time? There are two possible reasons. First, the proportion of self-citations in total citations is increasing over time (Hall et al, 2001). This might explain the increasing time trend of border and distance effects with self-citations since we know that self-citations exaggerate the magnitude of those effects. Second, the proportion of new knowledge flows is increasing over time. Since new knowledge faces larger barriers in diffusion, this will lead to upward slope of time trends of both border and distance effects.

4.3 Sources of Border Effects

We have shown that part of border effects come from self-citations. Of total citations in our sample, only 11% are self-citations, but they account for approximately 50% MSA and national border effects.[20] In other words, including self-citations approximately doubles those border effects.

The second source of overestimated aggregate border effects is aggregation bias. We find that there are at least three types of aggregation bias in the context of knowledge flows: geographic aggregation bias, age aggregation bias and category aggregation bias.

[19] In Jaffe, Trajtenberg and Henderson (1993), they argue that localization fades over time, but only very slowly.

[20] This proportion (11%) is consistent with the lower bound of the mean percentage of self-citations in the entire NBER database (Hall, Jaffe and Trajtenberg, 2001).

First, geographic aggregation bias substantially overestimated subnational border effects. The experiment is to decompose data only to state level and to compare the result with previous estimates. We find that the magnitude of state border effect is similar to the previous MSA border effect. However, if we further decompose data to the MSA level as in Table 4, we find that the state border effect almost vanishes as long as the MSA border is included. Also, if we estimate state border as the only subnational border using MSA level data as in specification (3) and (9) of Table 4, the magnitude of both subnational and national border effects becomes much smaller. This further suggests the existence of geographic aggregation bias for border effects in knowledge flows. Also, we have shown that more than 90% of subnational border effects come from metropolitan level rather than state. It implies that state border effect for knowledge flow is largely an artifact of geographic aggregation. This is consistent with the findings of trade flow in Hillberry and Hummels (2005), who argue that the state level home bias in trade flow is largely artifact of geographic aggregation.

Second, we find that decomposing data by different age group also reduces the size of border effects (see Table 5). The estimate of aggregate national border effect is around 1.14 to 2.69 larger than the estimates by age group, and the estimate of aggregate MSA border effect is around 1.22 to 2.27 larger than the disaggregated estimates. It is to some extent surprising since we have shown that new knowledge faces the largest barriers (border and distance effects) in diffusion. Hence we should expect that the magnitude of aggregate border effects is between the estimates from newest and oldest age groups. However, the aggregate border effect is always larger than the estimates in each age group, even the newest age group. It is hard to explain this phenomenon without age aggregation bias.

Third, decomposing data by category also helps to reduce border effects (see Table 6). This category aggregation bias might be related to some industrial "specialization" effect. Is that the case that the specialization matters rather than the true border matters? If we decompose the knowledge flows by category or by industry, can we eliminate the border effects? To answer this question, we need to look at the knowledge flows at the industry level. At a first glance, the rough category level result will give us some insights. In NBER Patent Citations Database, we have 6 rough categories: Chemical, Computers and Communications, Drugs and Medical, Electronics and Electricity, Mechanical, and Others. If border effects mostly stem from the specialization effect, then we should see a substantial decrease when we use decomposed data by category. We find that border effects do decrease, but not too much. Border effects are still there and significant. When we use the subsamples by category, the border effects are smaller. It means that some part of border effects come from the "specialization" effects. Once we split the sample by category, we alleviate some part of the border effects through ruling out the specialization effects. But the point is, specialization cannot explain all border effects. Also, specialization varies by industry. We prefer to call this type of bias "category aggregation bias" and it captures all bias due to the category or industry decomposition.

5 Robustness to Alternative Specifications

To see whether the time trend and the age profile of border and distance effects are robust, we examine several different specifications.

First, we verify the time trend of these effects. It is reasonable to take into account that there might be some interactions between time effect and age effect. For the whole sample, the time trend of border and distance effects are increasing. But if we only look at one particular subsample with similar age, does the time trend still hold? Figure 8, 9 and 10 illustrate the time trend of border and distance effects for each age group without self-citations. We find that border and distance indeed increase over time. Time trends are robust, even for different age group. However, the distance effect is more volatile in the upgrading trend. For very old knowledge flows (age greater than 20 years), distance effect is not significant. We also draw the time trends for different category without self-citations (see Figure 11, 12 and 13). The results show that for all 6 different categories, all border and distance effects are increasing over time. This again confirms the robust time trends.

Second, we want to examine whether the age profiles hold within each category. The results are noisy. Cat 6 (Others) still has decreasing border and distance effects with age. But other categories do not show the continuous decreasing age profiles. Some of the border and distance effects decrease with age first, but then start to increase in their very old life time periods. Also, some estimates for border and distance effects are not statistically significant. This implies that the category heterogeneity is huge in knowledge flows. Knowledge spillovers of different category or technological class behave very differently. The classification by category might be too broad to capture the industry level heterogeneity. Finer industry level data will be helpful to further the study. Controlling for technological difference between regions might be also helpful to examine this issue.

So far our results stem from the conventional estimation method in gravity literature without considering the zero flows. In the above results, we only include the positive citation flow. Hence, we only observe the knowledge flows with positive citation flow, i.e., c_{ij} is left censored at zero. But if a region doesn't make any, nor receive any citations, it also conveys an important information: the barriers of knowledge flows between these two regions are too high such that the barriers completely impede the knowledge flows. Hence it is better to take into account all zero flows to our basic framework. There are several ways to handle this problem. First, we can use a left-censored Tobit model. We find that Tobit estimates for aggregate border and distance effects are significant and they decrease with the age of knowledge. But Tobit estimates are biased due to a fundamental problem: if we use Tobit model, we assume that there might be some negative zero flows, just we cannot observe them and all observations are left-censored at zero. But in reality, we only have zero flows and positive flows. The quantity of knowledge flows is never negative. Hence, only using left-censored Tobit model is not the best choice to the question we study here. However, our aggregate level results are robust using Tobit estimates. The second method is developed by Helpman, Melitz, and Rubinstein (2007). They use two steps: first, estimate the probability of positive knowledge flows between each region pair; then, use predicted value to estimate the new gravity equation. To do this, we need to modify our simplest framework to include the zero citations between inventors in different locations. This has not been done in the current paper and it is one of the objectives of future work.

6 Conclusion

This paper employs a gravity framework to investigate the distance and border effects in knowledge spillovers, using evidence from patent citations panel data at metropolitan level within the U.S. and the 38 largest patent-cited countries outside the U.S. We present three key findings. First, we find strong subnational localization effects at the Metropolitan Statistical Area and state levels: more than 90% of intranational border effects stem from the metropolitan level rather than state. This contributes to the literature on subnational knowledge localization. Second, we characterize the age distribution of different types of knowledge spillovers and find that aggregate border and distance effects decrease with the age of cited patent. It implies the largest barriers in diffusion for new knowledge transmission, which is consistent with the intuition. However, over time, border and distance effects are interestingly increasing. We think that the increasing proportion of self-citations and new knowledge flows might explain this phenomenon. The age profiles and time trends of border and distance effects in knowledge spillovers are novel findings in the literature. Finally, we find that self-citations and aggregation bias are two sources of overestimated aggregate border effects of knowledge spillovers. Of total citations, only 11% are self-citations, but excluding self-citations approximately halves the MSA and national border effects. Decomposing data to finer geographic levels, by different age group and by different category substantially reduces the size of border effects. Among the three types of aggregation bias, geographic aggregation bias especially exaggerates the state border effect.

Appendix: Derivation of Gravity Equation of Knowledge Spillovers

We construct a simple gravity model of patent citations as exchange of ideas. We use patent citations, rather than patents themselves, as a proxy for knowledge products because (i)the number of citations constructs a flow variable we want while patent is usually treated as a stock variable in the previous literature;[21] (ii)citations carry information on the value of patents and concern inventors more than patents.[22]

Following the convention of gravity literature, in the benchmark model, we assume that: (1) Knowledge is differentiated by place of origin. Each region specializes in the production of a single knowledge product.[23] This assumption is widely used in the trade literature using gravity model, due to some simple observations, for example, Japanese rice is different from Thailand rice. It might be transplanted to knowledge flows, since it is not too hard to imagine that a citation to a German auto patent is not as same as a citation to a French auto patent. (2) All regions have the same tastes for the existing knowledge, i.e., identical, homothetic preferences, approximated by a CES function. (3) There exist barriers/frictions in knowledge flows. We have in mind information costs, design costs, and various legal and institutional costs, distance, organizational boundaries, language, etc. When inventors of new knowledge (patents) use the previous knowledge embodied in existing patents ("prior art"), they need to pay various forms of costs, for example, the translation of the foreign prior art, the examine fee for the patent examiners, and so on. (4) Markets for all knowledge products clear. (5) For each region, "inward" knowledge (citations made) from all source regions (including itself) is equal to its "outward" knowledge (citations received from all regions including itself). In other words, "exchange of knowledge" is balanced. This assumption is consistent with some observations in knowledge flows. The region citing more existing knowledge is also the region being cited more by others. For example, the U.S. is the largest knowledge both destination country and source country. The balance assumption is conventionally used in the gravity literature for trade flows, and the ratio of net outward knowledge to the sum of outward and inward of knowledge is even smaller than the net export to the sum of export and import of trade.[24] For example, the ratio of knowledge for the U.S. is -0.5% between 1975 and 2000 while the ratio of goods trade for the U.S. is -14.8% in 2000. Assumption (5) may be relaxed in future.

In the benchmark model, we only investigate one market: knowledge products market. In

[21]See Hall, Jaffe and Trajtenberg (2005), Bottazzi and Peri (2005). They consider the discount problem of the patent stock in their papers.

[22]See Trajtenberg (1990), Lanjouw and Schankerman (2004), Harhoff et al. (1999), Hall, Jaffe and Trajtenberg (2005).

[23]With this assumption we can suppress finer classifications of knowledge flows.

[24]I calculated the ratio of net outward knowledge to the sum of outward and inward of knowledge for main countries and found that most of them belong to [-20%, 10%].

this market, the individual inventor is both the consumer of existing knowledge products and the producer of new knowledge products. To keep things simple, in the current paper, we abstract from the heterogeneity of inventors and focus on the consumption behavior of knowledge products, since introducing the production behavior of knowledge needs to take into account other inputs (for example, the R&D expenditure and human capital) which we put into the future work.

When region i uses the previous knowledge, the citations occur, generating knowledge flows from region j to i. Through this way, representative inventors in region i "consume" these inward knowledge products from region j. Recall that we denote knowledge products by patent citations, then the problem for region i is to choose how many knowledge products to consume from each source region j, i.e., how many citations to make from each region j (citation c_{ij}) by maximizing its consumption of existing knowledge products in the world wide base:

$$\max_{c_{ij}} \quad (\textstyle\sum_j c_{ij}^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$$

subject to

$$\sum_j c_{ij} = y_i.$$

were σ is the elasticity of substitution between all knowledge products (patent citations); c_{ij} is the citation quantity from region j to region i; y_i is region i's total knowledge products which refer to the received citations (i.e., the total citations received by region i's patents). By assumption (5), y_i is equal to i's total outward knowledge (total citations j makes from all sources).

In the current model, the quantity of patent citation flows differ between the citing region and the cited region due to the existence of knowledge flow barriers that are not directly observable, and the main objective of the empirical work in this paper is to illustrate various patterns of these barriers and to identify them. If there are no barriers in knowledge flows, each region will get the same opportunity, based on their knowledge production capacity, to use the existing knowledge products in the world wide base. If the quantity of patent citations from region j (to an average destination region) is c_j, it will be $c_{ij} = c_j t_{ij}$ in region i when it arrives at region i. Here c_j denotes the cited region's supply quantity, net of knowledge flow frictions, c_{ij} is the real quantity from j to i and t_{ij} denotes the friction factor between j and i. Isoelastic demands imply that knowledge flows from j to i are given by,

$$c_{ij} = y_i(\frac{t_{ij}c_j}{Q_i})^{1-\sigma}$$

where Q_i is the knowledge flow quantity index of i, given by

$$Q_i = (\sum_j (t_{ij} c_j)^{1-\sigma})^{\frac{1}{1-\sigma}}$$

The general-equilibrium structure of the model imposes market clearance, which implies,

$$y_j = \sum_i c_{ij} = \sum_i (\frac{t_{ij} c_j}{Q_i})^{1-\sigma} y_i = c_j^{1-\sigma} \sum_i (\frac{t_{ij}}{Q_i})^{1-\sigma} y_i, (\forall j).$$

Proposition 1: If we further assume that knowledge flow frictions are symmetric, so that $t_{ij} = t_{ji}$, $\forall(i,j)$, then the supply quantities c_j's that are solution to previous equations are such that,

$$s_j \equiv \frac{y_j}{y_w} = (c_j Q_j)^{1-\sigma}$$

where the total quantity of knowledge production in the world is given by $y_w = \sum_j y_j$ and the knowledge production shares of region j is given by s_j.

Proof: If $\forall j$,$c_j^{1-\sigma} = s_j / Q_j^{1-\sigma}$, then,

$$\sum_i (\frac{t_{ij} c_j}{Q_i})^{1-\sigma} y_i = \sum_i \frac{y_j}{y_w Q_j^{1-\sigma}} (\frac{t_{ij}}{Q_i})^{1-\sigma} y_i = \frac{y_j}{Q_j^{1-\sigma}} \sum_i s_i (\frac{t_{ij}}{Q_i})^{1-\sigma} = \frac{y_j}{Q_j^{1-\sigma}} \sum_i (c_i t_{ij})^{1-\sigma}$$

by symmetry, $t_{ij} = t_{ji}$, so that,

$$\sum_i (\frac{t_{ij} c_j}{Q_i})^{1-\sigma} y_i = \frac{y_j}{Q_j^{1-\sigma}} \sum_i (t_{ji} c_i)^{1-\sigma} = \frac{y_j}{Q_j^{1-\sigma}} Q_j^{1-\sigma} = y_j.$$

Q.E.D.

With Proposition 1, we achieve a very useful simplification of gravity equation prediction for bilateral knowledge flows:

$$c_{ij} = \frac{y_j y_i}{y_w} (\frac{t_{ij}}{Q_j Q_i})^{1-\sigma}$$

with the quantity indices solution to,

$$Q_i^{1-\sigma} = \sum_j Q_j^{\sigma-1} s_j t_{ij}^{1-\sigma}, \forall i.$$

This provides an implicit solution to knowledge flow quantity indices as a function of all bilateral knowledge flow barriers and knowledge production shares. The quantity indices Q_j are referred to measures of "multilateral knowledge flow resistance" variables as they depend on all bilateral resistances t_{ij}.

This constructs our basic gravity model for knowledge flows. The gravity model tells us that bilateral knowledge flows, after controlling for size, depend on the bilateral knowledge flow frictions between i and j, relative to the product of their multilateral resistance indices.

The final step is to model the unobservable knowledge flow friction factor t_{ij}. We follow the convention in trade literature in hypothesizing that t_{ij} is a loglinear function of observables, bilateral distance d_{ij}, and whether there is a border b_{ij}:

$$lnt_{ij} = \tau_{ij} + \rho lnd_{ij} + \varepsilon_{ij} \tag{3}$$

where τ_{ij} is any other "border effect" associated with knowledge flows from region j to i. Generally t_{ij} is meant to include *all effects* limiting knowledge flows between i and j. Then we decompose τ_{ij} to subnational and national border indicators. $lnt_{ij} = B_{ij}^m + B_{ij}^s + B_{ij}^n + \rho lnd_{ij} + \varepsilon_{ij}$. Then following Anderson and Wincoop (2003) and Feenstra (2002), we get the theoretical gravity equation:

$$\ln(\frac{c_{ij}}{y_i y_j}) = k + (1-\sigma)\rho lnd_{ij} + (1-\sigma)B_{ij}^m + (1-\sigma)B_{ij}^s + (1-\sigma)B_{ij}^n + \ln(Q_i)^{\sigma-1} + \ln(Q_j)^{\sigma-1} + (1-\sigma)\varepsilon_{ij}$$

where k is a constant. Then we can rewrite the above equation as

$$\ln(\frac{c_{ij}}{y_i y_j}) = k + \alpha lnd_{ij} + \beta_1 B_{ij}^m + \beta_2 B_{ij}^s + \beta_3 B_{ij}^n + \ln(Q_i)^{\sigma-1} + \ln(Q_j)^{\sigma-1} + (1-\sigma)\varepsilon_{ij}$$

This is our theoretical gravity equation of knowledge flows.

References

Adams, J. D. (2002): "Comparative Localization of Academic and Industrial Spillovers," *Journal of Economic Geography*, 2:3, 253–278.

Alcácer, J., and M. Gittelman (2006): "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations," *The Review of Economics and Statistics*, 88(4), 774–779.

Almeida, P. (1996): "Knowledge Sourcing by Foreign Multinationals: Patent Citation Analysis in the U.S. Semiconductor Industry," *Strategic Management Journal*, 17, 155–165.

Almeida, P., and B. Kogut (1999): "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, 45(7), 905–917.

Anderson, J. E., and E. V. Wincoop (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–192.

——— (2004): "Trade Costs," *Journal of Economic Literature*, XLII, 691–751.

Audretsch, D. B., and M. P. Feldman (2004): "Knowledge Spillovers and the Geography of Innovation," *Chapter 61, Handbook of Regional and Urban Economics*, 4, 2713–2739.

Basberg, B. L. (1987): "Patents and the Measurement of Technological Change: A Survey of the Literature," *Research Policy*, 16, 131–141.

Branstetter, L. G. (1996): "Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the Us and Japan," *Journal of International Economics*.

——— (2004): "Is Foreign Direct Investment a Channel of Knowledge Spillovers? Evidence from Japan's FDI in the United States," *NBER Working Paper*.

Branstetter, L. G., and M. Sakakibara (2002): "When Do Research Consortia Work Well and Why? Evidence from Japanese," *American Economic Review*, 92(1), 143–159.

Caballero, R., and A. Jaffe (1993): "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," *NBER Macroeconomics Annual*, Cambridge. MA: MIT Press.

Cairncross, F. (1997): "The Death of Distance: How the Communications Revolution Will Change Our Lives," Harvard: Harvard Business School Press.

Carpenter, M. P., and F. Narin (1983): "Validation Study: Patent Citations as Indicators of Science and Foreign Dependence," *World Patent Information*, 5(3), 180–185.

Coe, D. T., and E. Helpman (1995): "International R&D spillovers," *European Economic Review*, 39(5), 859–887.

Combes, P.-P. (2008): "gravity models," *The New Palgrave Dictionary of Economics, Eds. Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan*, The New Palgrave Dictionary of Economics Online, Palgrave Macmillan., DOI:10.1057/9780230226203.0670.

Coyle, D. (1997): "Weightless World," Princeton: Princeton University Press.

Eaton, J., and S. Kortum (1996): "Trade in Ideas: Patenting and Productivity in the OECD," *Jounral of International Economics*, 40(3-4), 251–278.

——— (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–1779.

Feenstra, R. C. (2002): "Border Effects and the Gravity Equation: Consistent Methods for Estimation," *Scottish Journal of Political Economy*, 49, 491–506.

——— (2003): "Advanced International Trade," Princeton University Press.

Feldman, M. P., and D. B. Audretsch (1999): "Innovation in Cities: Science-based Diversity, Specialization and Localized Competition," *European Economic Review*, 43, 409–429.

Friedman, T. (2005): "The World is Flat," New York: Farrar, Strauss and Giroux.

Griffith, R., S. Lee, and J. V. Reenen (2007): "Is Distance Dying at Last? Falling Home Bias in Fixed Effects Models of Patent Citations," *NBER Working Paper No. 13338.*

Griliches, Z. (1990): "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28, 1661–1707.

Griliches, Z., and F. R. Lichtenberg (1984): "Interindustry Technology Flows and Productivity Growth: A Reexamination," *Review of Economics and Statistics*, LXVI(2), 324–329.

Griliches, Z., A. Pakes, and B. H. Hall (1987): "The Value of Patents as Indicators of Inventive Activity," *Economic Policy and Technological Performance*, P. Dasgupta and P. Stoneman (eds). Cambridge, England: Cambridge University Press, 97–124.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2001): "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," *NBER Working Paper No. 8498.*

——— (2005): "Market Value and Patent Citations," *Rand Journal of Economics*, 36.

Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel (1999): "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 81(3), 511–515.

Helpman, E., M. Melitz, and Y. Rubinstein (2007): "Estimating Trade Flows: Trading Partners and Trading Volumes," *NBER Working Paper No. 12927.*

Henderson, R., A. B. Jaffe, and M. Trajtenberg (1998): "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting 1965-1988," *Review of Economics and Statistics*, pp. 119–127.

Hillberry, R., and D. Hummels (2005): "Trade Responses to Geographic Frictions: a Decomposition Using Micro-data," *NBER Working Paper No. 11339.*

Hu, A. G. Z., and A. B. Jaffe (2001): "Patent Citations and International Knowledge Flow: The Cases of Korea and Taiwan," *NBER Working Paper No. W8528.*

Hummels, D. (1999): "Toward a Geography of Trade Costs," Mimeo, Krannert School of Management, Purdue University.

Jaffe, A. B. (1986): "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, 76, 984–1001.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 577–598.

Keller, W. (1998): "Are international R&D spillovers trade-related? Analyzing spillovers among randomly matched trade partners," *European Economic Review*, 42(8), 1469–1481.

Lee, G. (2006): "The Effectiveness of International Knowledge Spillover Channels," *European Economic Review*, 50, 2075–2088.

McCallum, J. (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review*, 85(3), 615–623.

Monjon, S., and P. Waelbrock (2003): "Assessing Spillovers from Universities to Firms: Evidence from French Firm-Level Data," *International Journal of Industrial Organization*, 21(9), 1255–1270.

Navaretti, G. B., and I. Soloaga (2002): "Weightless Machines and Costless Knowledge: An Empirical Analysis of Trade and Technology Diffusion," *CEPR Discussion Paper.*

Peri, G. (2003): "Knowledge Flows, R&D Spillovers, and Innovation," *Mannheim, Germany: ZEW Discussion Paper.*

——— (2005): "Determinants of Knowledge Flows and Their Effect on Innovation," *The Review of Economics and Statistics*, 87(2), 308–322.

Powell, J. L. (1984): "Least Absolute Deviations Estimation for the Censored Regression Model," *Journal of Econometric*, 25, 303–325.

——— (1986): "Symmetric Trimmed Least Squares Estimation for Tobit Models," *Econometrica*, 54, 1435–1460.

Shane, H., and M. Klock (1997): "The Relation between Patent Citations and Tobin's Q in the Semiconductor Industry," *Review of Quantitative Finance and Accounting*, 9(2), 131–146.

Thompson, P. (2006): "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations," *The Review of Economics and Statistics*, 88(2), 383–388.

Thompson, P., and M. Fox-Kean (2005): "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment," *American Economic Review*, 95(1), 450–460.

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

Zucker, L. G., M. R. Darby, and J. Armstrong (1998): "Intellectual Capital and the Firm: The Technology of Geographically Localized Spillovers," *Economic Inquiry*, 36, 65–86.

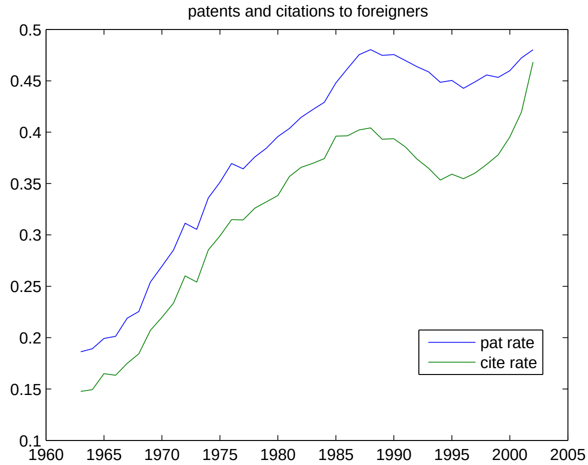

Figure 1. The share of patents granted to foreigners and citations generated by foreigners.

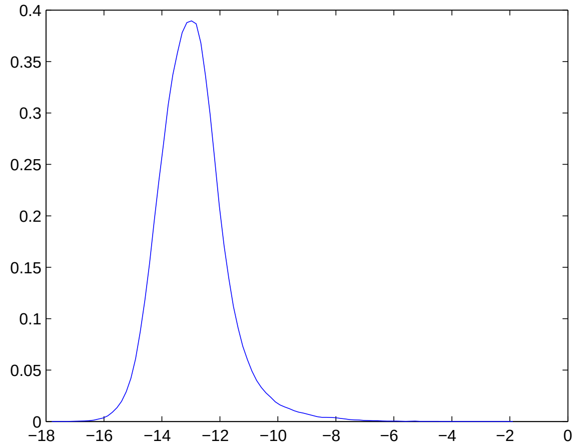

Figure 2. The distribution of $ln(\frac{c_{ij}}{y_i y_j})$.

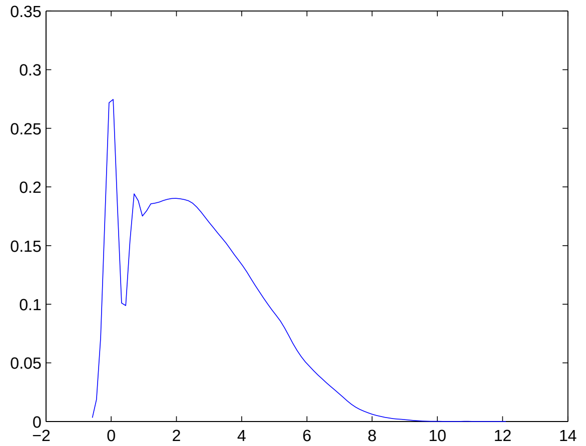

Figure 3. The distribution of $ln(c_{ij})$.

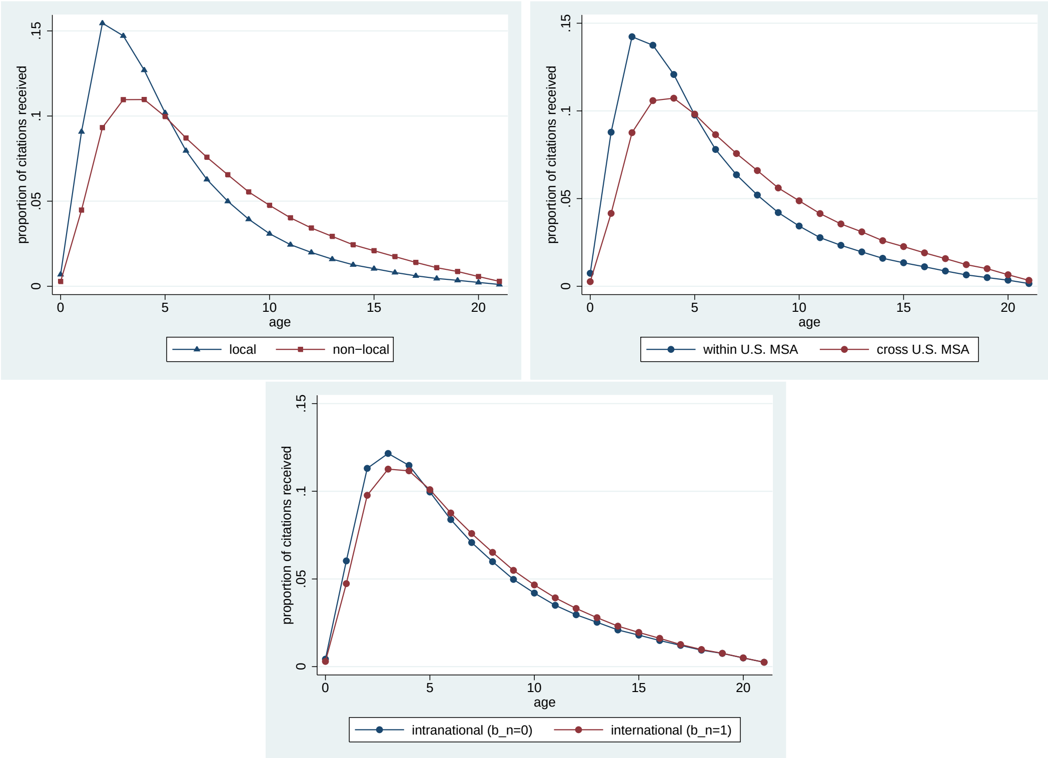

Figure 4. The age distribution of knowledge diffusion (with self-citation).

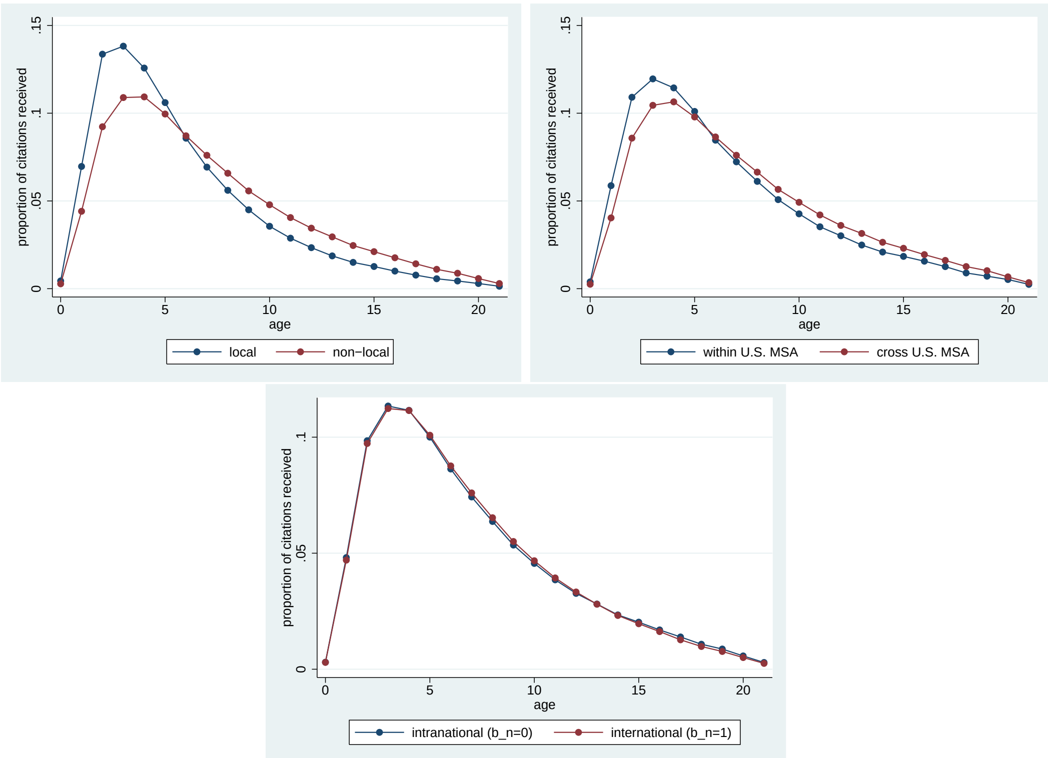

Figure 5. The age distribution of knowledge diffusion (without self-citation).

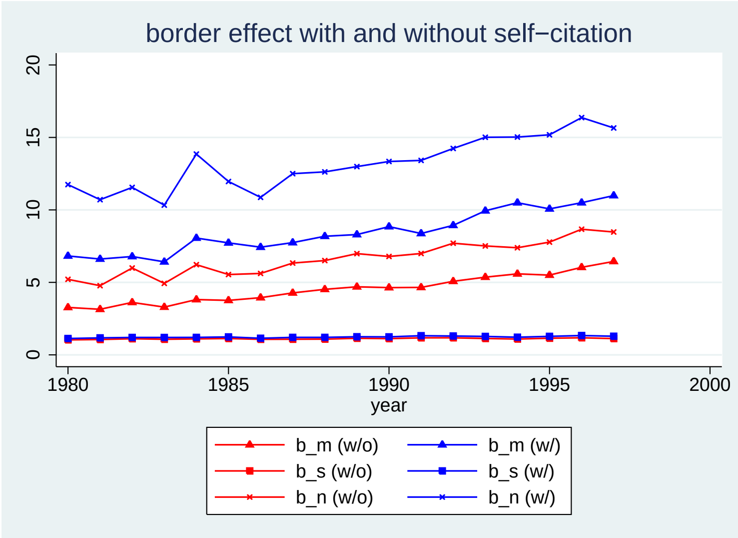

Figure 6. Time trends of aggregate border effects (with and without self-citation).

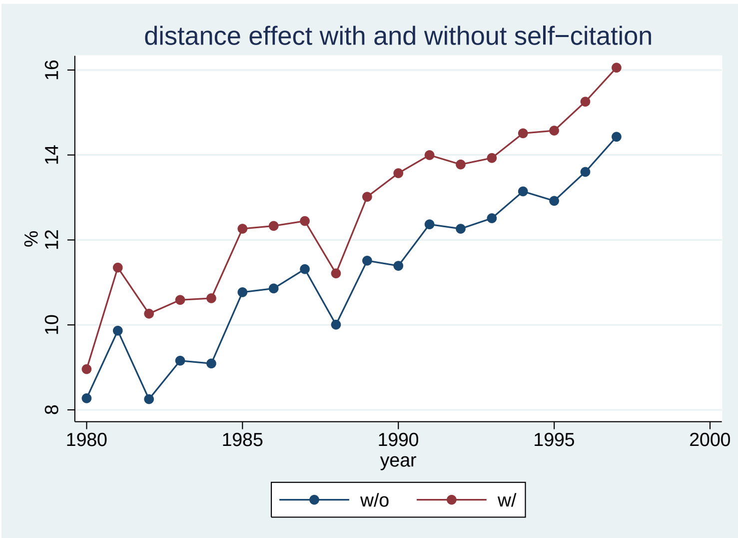

Figure 7. Time trends of aggregate distance effects (with and without self-citation).

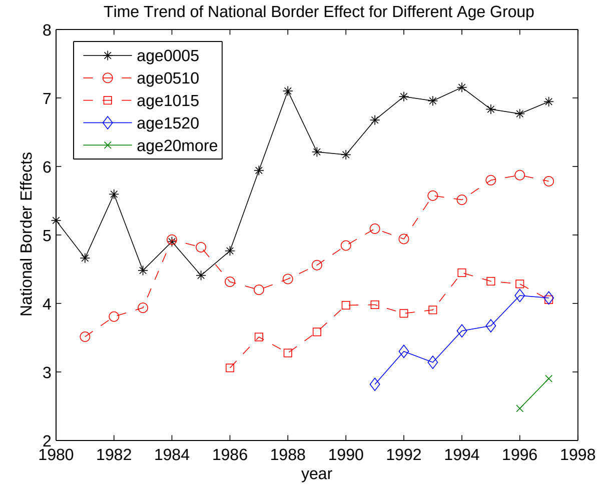

Figure 8. Time Trend of National Border Effects for Different Age Group.

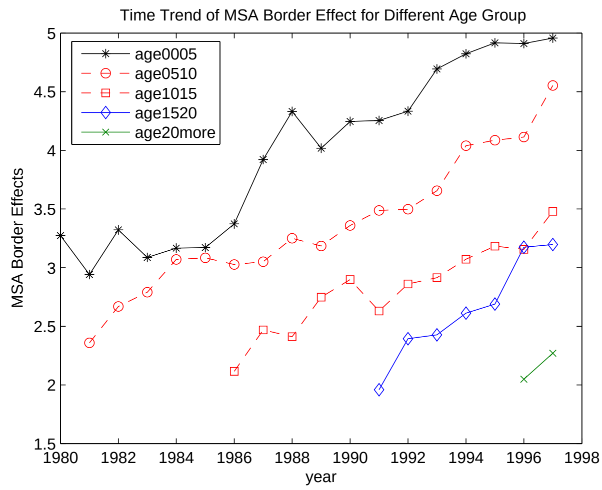

Figure 9. Time Trend of MSA Border Effects for Different Age Group.

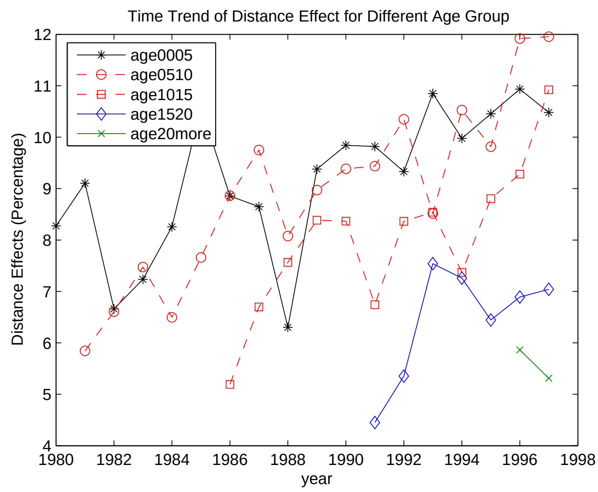

Figure 10. Time Trend of Distance Effects for Different Age Group.

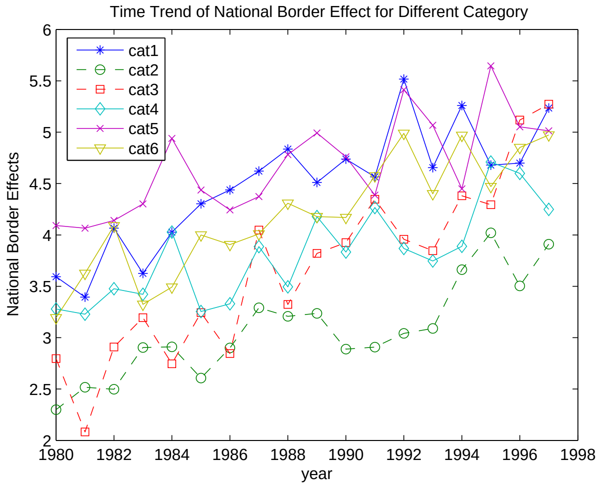

Figure 11. Time Trend of National Border Effects for Different Category.

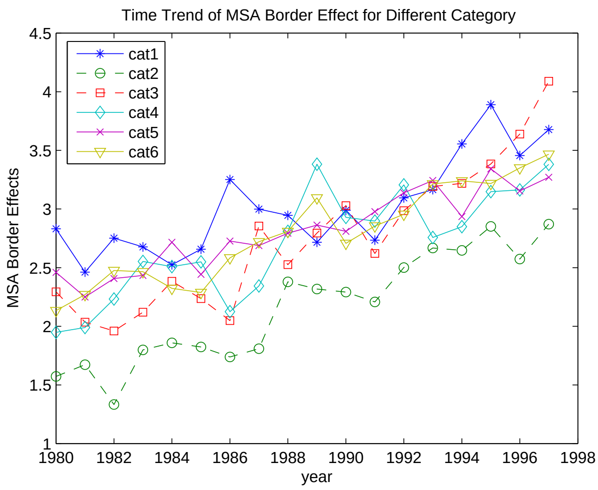

Figure 12. Time Trend of MSA Border Effects for Different Category.

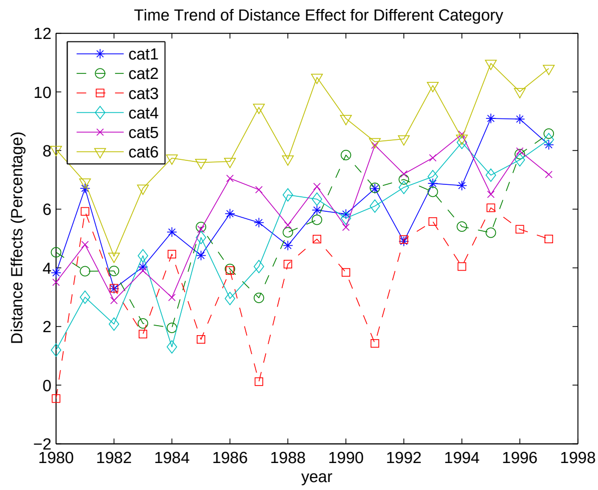

Figure 13. Time Trend of Distance Effects for Different Category.

Table 1. Rank of Knowledge Production Capacity (1980-1997)

Rank	Economy	Yearly received citations (with self-citation)	Rank	Economy	Yearly received citations (without self-citation)
1	U.S.	218531	1	U.S.	194310
2	JAPAN	70553	2	JAPAN	59932
3	GERMANY	26024	3	GERMANY	23095
4	UNITED KINGDOM	11586	4	UNITED KINGDOM	10748
5	FRANCE	9782	5	FRANCE	9031
6	CANADA	6392	6	CANADA	6044
7	SWITZERLAND	4867	7	SWITZERLAND	4253
8	ITALY	3386	8	ITALY	3161
9	NETHERLANDS	3312	9	SWEDEN	3074
10	SWEDEN	3210	10	NETHERLANDS	2935
11	AUSTRALIA	1333	11	AUSTRALIA	1296
12	TAIWAN	1304	12	TAIWAN	1269
13	BELGIUM	1198	13	BELGIUM	1039
14	AUSTRIA	1053	14	AUSTRIA	937
15	ISRAEL	951	15	ISRAEL	898
16	FINLAND	703	16	FINLAND	648
17	DENMARK	664	17	DENMARK	622
18	RUSSIA	621	18	RUSSIA	618
19	SOUTH KOREA	596	19	SOUTH KOREA	534
20	SOUTH AFRICA	353	20	SOUTH AFRICA	342

Table 2. Representative Regions (1980-1997)

Panel A: Representative High-Cited Regions		
Region	Yearly received citations (with self-citation)	Yearly received citations (without self-citation)
Japan	70553	59932
Germany	26024	23095
New York-Northern New Jersey-Long Island, NY-NJ-CT-PA (U.S.)	23630	21058
San Francisco-Oakland-San Jose, CA (U.S.)	16548	14838
Los Angeles-Riverside-Orange County, CA (U.S.)	13427	12619
Chicago-Gary-Kenosha, IL-IN-WI (U.S.)	12017	10705
United Kingdom	11586	10748
Boston-Worcester-Lawrence, MA-NH-ME-CT (U.S.)	9950	9193
France	9782	9031
Philadelphia-Wilmington-Atlantic City, PA-NJ-DE-MD (U.S.)	8675	7269
Panel B: Representative Low-Cited Regions		
Region	Yearly received citations (with self-citation)	Yearly received citations (without self-citation)
Eugene-Springfield, OR (U.S.)	8	7
Turkey	7	7
Laredo, TX (U.S.)	7	7
Grand Forks, ND-MN (U.S.)	6	6
Iceland	5	5
Anniston, AL (U.S.)	5	5
Jacksonville, NC (U.S.)	4	4

Table 3. Basic Estimation Results for Aggregate Border and Distance Effects

Specification:	(1)	(2)	(3)	(4)	(5)	(6)
lnd_{ij}	-0.131**	-0.154**	-0.211**	-0.167**	-0.198**	-0.259**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
B_{ij}^m	-2.134**	-2.245**		-2.431**	-2.574**	
	(0.014)	(0.013)		(0.021)	(0.020)	
B_{ij}^s	-0.224**		-0.655**	-0.289**		-0.780**
	(0.009)		(0.009)	(0.014)		(0.013)
B_{ij}^n	-2.589**	-2.433**	-0.858**	-2.897**	-2.697**	-0.926**
	(0.018)	(0.017)	(0.015)	(0.027)	(0.026)	(0.022)
B_{ij}^m effect	8.449**	9.440**		11.375**	13.120**	
	(0.119)	(0.126)		(0.240)	(0.263)	
B_{ij}^s effect	1.252**		1.925**	1.335**		2.181**
	(0.011)		(0.017)	(0.018)		(0.028)
B_{ij}^n effect	13.316**	11.390**	2.360**	18.126**	14.829**	2.525**
	(0.243)	(0.196)	(0.034)	(0.498)	(0.383)	(0.055)
Citing-region fixed effects	yes	yes	yes	yes	yes	yes
Cited-region fixed effects	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes			
constant	yes	yes	yes	yes	yes	yes
No. of observations (ij,t)	473294	473294	473294	473294	473294	473294
F-statistics	1826	1825	1714	458	458	428
Adjusted R^2	0.74	0.74	0.73	0.41	0.41	0.39

Notes: ** Significant at 1% level.

Table 4. Aggregate Border and Distance Effects (with and without self-citations)

Specification:	(1)	(2)	(3)	(7)	(8)	(9)
	With self-citation			Without self-citation		
lnd_{ij}	-0.131**	-0.154**	-0.211**	-0.116**	-0.128**	-0.167**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
B_{ij}^m	-2.134**	-2.245**		-1.509**	-1.573**	
	(0.014)	(0.013)		(0.015)	(0.014)	
B_{ij}^s	-0.224**		-0.655**	-0.124**		-0.433**
	(0.009)		(0.009)	(0.009)		(0.009)
B_{ij}^n	-2.589**	-2.433**	-0.858**	-1.903**	-1.821**	-0.695**
	(0.018)	(0.017)	(0.015)	(0.019)	(0.018)	(0.015)
B_{ij}^m effect	8.449**	9.440**		4.524**	4.823**	
	(0.119)	(0.126)		(0.067)	(0.067)	
B_{ij}^s effect	1.252**		1.925**	1.132**		1.542**
	(0.011)		(0.017)	(0.011)		(0.014)
B_{ij}^n effect	13.316**	11.390**	2.360**	6.707**	6.178**	2.003**
	(0.243)	(0.196)	(0.034)	(0.126)	(0.109)	(0.029)
Citing-region fixed effects	yes	yes	yes	yes	yes	yes
Cited-region fixed effects	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes
constant	yes	yes	yes	yes	yes	yes
No. of observations (ij,t)	473294	473294	473294	467205	467205	467205
F-statistics	1826	1825	1714	1721	1723	1672
Adjusted R^2	0.74	0.74	0.73	0.73	0.73	0.72

Notes: ** Significant at 1% level.

Table 5. Estimates by Age of Knowledge (without Self-citation)

Specification:	whole sample	age [0,5)	age [5,10)	age [10,15)	age [15,20)	age [20,more)
lnd_{ij}	-0.116**	-0.092**	-0.091**	-0.079**	-0.065**	-0.059*
	(0.002)	(0.003)	(0.004)	(0.005)	(0.008)	(0.028)
B_{ij}^m	-1.509**	-1.312**	-1.167**	-0.991**	-0.866**	-0.691**
	(0.015)	(0.020)	(0.024)	(0.034)	(0.057)	(0.192)
B_{ij}^s	-0.124**	-0.108**	-0.095**	-0.069**	-0.071†	0.066
	(0.009)	(0.014)	(0.016)	(0.023)	(0.039)	(0.138)
B_{ij}^n	-1.903**	-1.769**	-1.530**	-1.315**	-1.191**	-0.913**
	(0.019)	(0.026)	(0.030)	(0.043)	(0.070)	(0.229)
MSA border effect	4.524**	3.713**	3.214**	2.694**	2.379**	1.996**
	(0.067)	(0.074)	(0.077)	(0.092)	(0.136)	(0.382)
state border effect	1.132**	1.114**	1.099**	1.071**	1.074†	1.068
	(0.011)	(0.016)	(0.018)	(0.025)	(0.041)	(0.148)
national border effect	6.707**	5.863**	4.618**	3.726**	3.289**	2.492**
	(0.126)	(0.151)	(0.140)	(0.159)	(0.229)	(0.570)
Citing-region effect	yes	yes	yes	yes	yes	yes
Cited-region effect	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes
constant	yes	yes	yes	yes	yes	yes
No. of observations (ij,t)	467205	283980	285081	169010	83960	14258
F-statistics	1721	824	710	399	232	46
Adjusted R^2	0.73	0.68	0.65	0.63	0.66	0.69

Notes: ** Significant at 1% level. * Significant at 5% level. † Significant at 10% level.

Table 6. Border and Distance Effects by Category (without Self-citation)

	Whole sample	Cat 1 Chemical	Cat 2 C.&C.	Cat 3 D.&M.	Cat 4 E.&E.	Cat 5 Mechanical	Cat 6 Others
lnd_{ij}	-0.116**	-0.055*	-0.056**	-0.020	-0.047**	-0.057**	-0.084**
	(0.002)	(0.009)	(0.015)	(0.014)	(0.009)	(0.006)	(0.004)
B_{ij}^m	-1.509**	-0.843**	-0.425**	-0.527**	-0.771**	-0.903**	-0.901**
	(0.015)	(0.057)	(0.100)	(0.088)	(0.061)	(0.035)	(0.028)
B_{ij}^s	-0.124**	-0.143**	-0.047	-0.096	-0.091*	-0.094**	-0.090**
	(0.009)	(0.044)	(0.074)	(0.070)	(0.046)	(0.026)	(0.020)
B_{ij}^n	-1.903**	-1.323**	-0.849**	-0.876**	-1.152**	-1.421**	-1.330**
	(0.019)	(0.070)	(0.117)	(0.103)	(0.073)	(0.044)	(0.035)
B_{ij}^m Effect	4.524**	2.323**	1.530**	1.693**	2.162**	2.467**	2.461**
	(0.067)	(0.131)	(0.153)	(0.149)	(0.132)	(0.086)	(0.068)
B_{ij}^s Effect	1.132**	1.154**	1.048	1.101	1.096*	1.098**	1.094**
	(0.011)	(0.050)	(0.077)	(0.077)	(0.050)	(0.028)	(0.021)
B_{ij}^n Effect	6.707**	3.753**	2.338**	2.402**	3.164**	4.142**	3.781**
	(0.126)	(0.263)	(0.273)	(0.247)	(0.231)	(0.184)	(0.132)
Citing effect	yes	yes	yes	yes	yes	yes	yes
Cited effect	yes	yes	yes	yes	yes	yes	yes
Year effect	yes	yes	yes	yes	yes	yes	yes
constant	yes	yes	yes	yes	yes	yes	yes
No. of observations (ij,t)	467205	128987	84978	94177	123681	169061	222546
F-statistics	1721	214	177	174	233	413	554
Adjusted R^2	0.73	0.55	0.59	0.57	0.58	0.64	0.65

Notes: ** Significant at 1% level. * Significant at 5% level. † Significant at 10% level.