INTERNATIONAL KNOWLEDGE FLOWS: EVIDENCE FROM PATENT CITATIONS

Adam B. Jaffe
Manuel Trajtenberg

Working Paper **6507**

NBER WORKING PAPER SERIES

INTERNATIONAL KNOWLEDGE FLOWS: EVIDENCE FROM PATENT CITATIONS

Adam B. Jaffe
Manuel Trajtenberg

Working Paper 6507
http://www.nber.org/papers/w6507

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 1998

We acknowledge with gratitude financial support from the National Science Foundation, Grant No. SBR-9320973 to the National Bureau of Economic Research, and No. SBR-9413099, and from the Israeli Science Foundation, Grant 874/96. We appreciate comments from seminar participants at Brandeis University, Harvard University, NBER Science and Technology Policy Research Group, NBER Productivity Summer Institute, and Northwestern University. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

International Knowledge Flows: Evidence
from Patent Citations
Adam B. Jaffe and Manuel Trajtenberg
NBER Working Paper No. 6507
April 1998
JEL No. O31

ABSTRACT

This paper explores the patterns of citations among patents taken out by inventors in the U.S., the U.K., France, Germany and Japan. We find (1) patents assigned to the same firm are more likely to cite each other, and come sooner than other citations; (2) patents in the same patent class are approximately 100 times as likely to cite each other as patents from different patent classes, but there is not a strong time pattern to this effect; (3) patents whose inventors reside in the same country are typically 30 to 80% more likely to cite each other than inventors from other countries, and these citations come sooner; and (4) there are clear country-specific citation tendencies; e.g., Japanese citations typically come sooner than those of other countries.

Adam B. Jaffe
Department of Economics
Mailstop 021
Brandeis University
Waltham, MA 02254-9110
and NBER
ajaffe@nber.org

Manuel Trajtenberg
Eitan Berglas School of Economics
Tel-Aviv University
Tel-Aviv 69978
ISRAEL
and NBER
manuel@ccsg.tau.ac.il

I. INTRODUCTION

The rate at which knowledge diffuses outward from the institutional setting and geographic location in which it is created has important implications for the modeling of technological change and economic growth, and for science and technology policy. Models of endogenous economic growth, such as Romer (1990) or Grossman and Helpman (1991), typically treat knowledge as completely diffused within an economy, but implicitly or explicitly assume that knowledge does not diffuse across economies. In the policy arena, ultimate economic benefits are increasingly seen as the primary policy motivation for public support of scientific research. Obviously, the economic benefits to the U.S. economy of domestic research depend on the fruits of that research being more easily or more quickly harvested by domestic firms than by foreign firms. Thus for both modeling and policy-making purposes it is crucial to understand the institutional, geographic and temporal dimensions of the spread of newly created knowledge.

There is an existing empirical literature on international technology flows. Much of this literature focuses on what might be described as "technology" diffusion rather than knowledge diffusion. For example, Teece (1977) discusses the difficulties that a multinational firm has in applying technology developed in one country to its operations overseas. Park (1995) and Coe and Helpman (1995) examine the impact on a country's productivity growth of the trade-weighted R&D of other countries. Generally, a positive effect is found, which can be interpreted as reduced-form evidence of knowledge spillovers across international boundaries. While the mechanism for such spillovers is not identified, it seems reasonable that many forms of communication and information transfer would be correlated with bilateral trade flows. In these analyses, however, it is difficult to distinguish the effect of "pure" knowledge flows from the effect of technology flows embodied in advanced capital goods sold from one country to another. This distinction is crucial. Knowledge is inherently nonrival in its use, and hence its creation and diffusion are likely to lead to spillovers and increasing returns; it is this nonrival property of knowledge that is at the theoretical heart of models that produce endogenous growth from research. But to the extent that the knowledge or technology flow is embodied in a purchased piece of equipment, it may not produce a spillover, or, if it does, the spillover may take the form of a pricing or pecuniary externality rather than a technological one (Griliches, 1979).

Knowledge spillovers are much harder to measure than technology transfer, precisely because they tend to be disembodied. In previous work (Jaffe and Trajtenberg, 1996; Jaffe, Henderson and Trajtenberg, 1993), we have looked at citations made by patents to previous patents as a "window" on the process of knowledge flow. Jaffe, Henderson and Trajtenberg, 1993, showed that patent citations do appear to be somewhat localized geographically, implying that a region or country does utilize knowledge created within it somewhat more readily than do more remote regions. In Jaffe and Trajtenberg, 1996, we went further, looking in detail at citations from other countries' patents to those of the U.S. We showed there that there is a clear time path to the diffusion of knowledge, in which domestic inventors' citation probabilities are particularly high in the early years after an invention is made.

While this previous work indicates the usefulness of patent citations for exploring knowledge flows, it also highlights the need for careful attention to the details of the patenting and citation processes. In particular, changes in citation practices, truncation biases, technology field effects, and the presence of large numbers of "self-citations" must all be taken into account in using citation data to examine knowledge flows.

We have three goals in this paper. First, we demonstrate how an econometric model can be used to make citations a potentially useful measure of knowledge flows, by controlling for the effects of truncation, changes in citation patterns, and technology field effects. Second, we explore for the first time the citation patterns among all combinations of the G-5 countries, the U.S., Great Britain, France, Germany and Japan. This gives a much richer picture of the geographic dimension of citation diffusion, by examining the extent and speed of diffusion of citations within and among all combinations of these countries. This permits us to estimate the extent and nature of "localization" of citations within each of these countries, to examine differences among the countries in their apparent absorption of foreign technology, and to identify some interesting pairwise interactions. Finally, we add the dimensions of "institutional localization" and "technological localization" to the modeling, and examine the interactions between localization in these dimensions and in geography.

II. KNOWLEDGE FLOWS AND PATENT CITATIONS

Consider a researcher or inventor working on a given technological problem at a given time in a given geographic location and institutional setting. This inventor might find it easier, cheaper or faster to solve her technological problem by virtue of access to knowledge created earlier by other inventors and researchers. For linguistic color and convenience, call the invention that is facilitated by some earlier piece of research the "descendant" and the earlier work that contributed to it the "antecedent." The question we want to ask is: how is the probability that a given descendant will benefit from a specific antecedent affected by the time, geographic location, institutional setting and technological nature of each, and by the relationship between the two along each of these dimensions. In particular, we are interested in the extent of "localization" in geography, institutional setting and technology space, and how localization interacts with time. That is, is a descendant more likely to benefit from an antecedent that is nearby geographically, comes from within the same institution, and is technologically similar, and does this increased likelihood of benefiting from nearby antecedents vary with the length of time elapsed.

Our expectation is that knowledge follows a diffusion process through geographic, institutional and technological spaces. Thus, researchers that are "nearby" along each of these dimensions would be particularly likely to benefit disproportionately in the time period immediately after the antecedent innovation occurs. We expect, however, that this "localization effect" will tend to fade over time, so that eventually the probability of an antecedent benefiting a remote descendant may be no lower than the probability of benefiting one nearby.

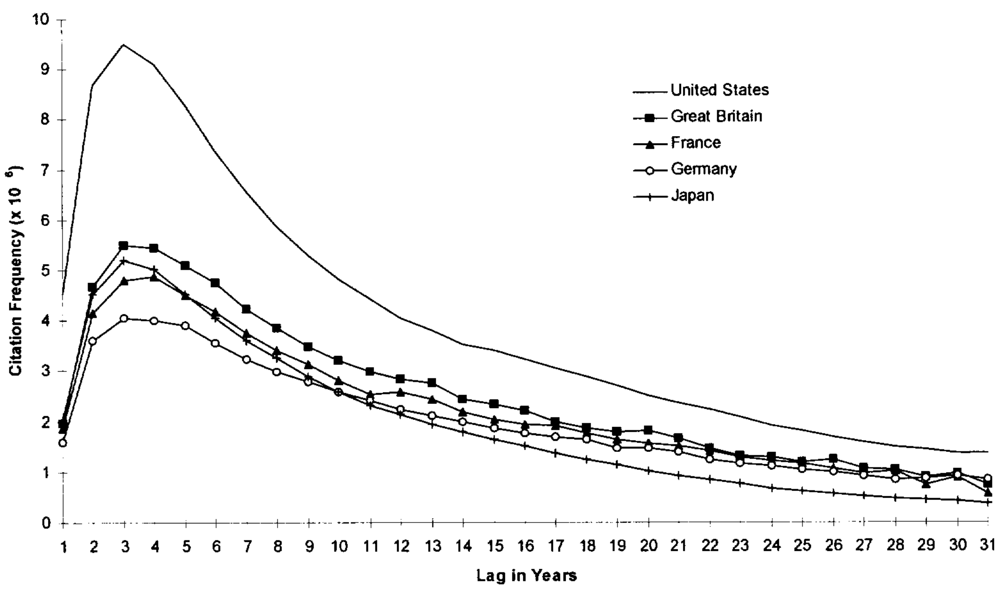
Thus localization and the fading of localization are phenomena that derive from the *relationship between* two inventions or inventors. But these relational phenomena are intrinsically tied up with the attributes of the antecedent and descendant themselves. A particular inventor (or group of inventors) may just be good at picking up and implementing others' ideas quickly, and others may be good at disseminating or spreading the implications of their research, or may produce research which is systematically more "fruitful" in stimulating others. The probability that a particular group will benefit from some other group (and the changes over time of this probability) will therefore be determined jointly by the properties of each group, and the properties of the relationship between the groups.

In addition to diffusing outward over time, bits of knowledge also become obsolete. Thus, though the probability that a given inventor will know of a given antecedent increases as the time lag between them grows, the probability that the antecedent will actually be helpful declines, on average. The combination of diffusion and obsolescence processes may cause the probability of using a given antecedent to first rise and then fall with elapsed time.

Our maintained assumption is that patents are a proxy for "bits of knowledge" and patent citations are a proxy for a given bit of knowledge being useful in the development of a descendant bit. This permits us to use the probability of citation as a proxy for the probability of useful knowledge flow, and empirical citation frequencies as a measure of that probability. Of course, the frequency with which generation of knowledge bits leads to a patent (the "propensity to patent") varies over time and space, as does the frequency with which use of earlier knowledge produces a citation (the "propensity to cite"). These variations in the correspondence between the data and the underlying constructs of interest create problems of interpretation that must be dealt with via a combination of multiple measurements and identifying assumptions.

The nature of these issues can be seen in Figure 1, which plots empirical citation frequencies from other countries to the U.S., as a function of the time lag between the citing and cited patents. The citation frequency is calculated as the total number of citations divided by the product of the number of potentially citing and number of potentially cited patents. For example, Japanese inventors took out about 22 thousand patents in 1993. U.S. inventors took out about 36 thousand patents in 1969. A total of about 800 citations from 1993-Japanese patents were made to 1969-U.S. patents. Hence the estimated citation frequency for this combination is about $1x10^{-6}$ (800/(22000*36000)). The citation frequencies plotted in Figure 1 are averages for all combinations with a given lag for which we have data, e.g., the calculated frequency at lag 30 derives from citations from 1993 to 1963 (our earliest data year) and 1994 (our last data year) to 1964. We interpret the citation frequency as an estimate of the probability that a randomly drawn patent in the citing group will cite a randomly drawn patent in the cited group.

Figure 1
Raw Citation Frequencies to U.S.-Invented Patents, by Citing Country



Important features of these data are immediately visible in Figure 1. First, the citation frequency rises rapidly in the first few years after the cited patent, then peaks and declines slowly over time. A significant number of citations are still being received many years after initial grant. Second, a U.S.-invented patent is much more likely to be cited by a U.S.-invented patent than it is by a foreign-invented patent. Finally, even putting aside the domestic/foreign distinction, there are noticeable differences in the citation frequencies across citing countries. For example, the likelihood of a random U.K.-invented patent citing a U.S.-invented patent is about 40% higher than the likelihood of a random German-invented patent citing a U.S.-invented patent.

Although all of these qualitative features illustrated in Figure 1 are in some sense "real," raw citation frequencies are afflicted by numerous theoretical and actual biases that make their interpretation dangerous. First, the observation of citations is always subject to truncation bias. Since we can observe only the citations already granted, we can see citations at long lags only for citations *from* very recent cohorts *to* very old cohorts. The significance of the truncation problem is greatly exacerbated by the fact that the number of citations *made* per patent has been rising significantly in the last two decades (Caballero and Jaffe, 1993). Thus the observations at long lags in Figure 1 all come from patents granted when relatively many citations were made (e.g., citations from 1993 patents to 1963 patents), whereas the observations at short lags are mixtures of many different cited cohort/citing cohort combinations (e.g., for lag=5 we have 1977 to 1972, 1987 to 1982, and so forth). We will see below that controlling for these interacting time effects yields predicted probabilities for long lags considerably lower than those shown in Figure 1.

In addition to artifacts of the citation process, the numbers in Figure 1 contain effects operating along the institutional and technological dimensions that interact non-randomly with geography. Not surprisingly, the probability that an inventor will cite another inventor employed by the same firm is much higher than the probability of citing a random inventor employed elsewhere. And, inventors employed by the same firm are more likely to live in the same country than random inventors employed by different firms. Hence the higher citation frequency for U.S. to U.S. than for Japan to U.S. is partly due to a higher citation frequency within firms, combined with a geographic localization of employees within firms. While for some purposes it might be appropriate to include this "firm self-citation" effect within what we call the geographic localization effect, for other purposes we may want to separate the two. Similarly, though ultimately less important empirically, an inventor is much more likely to cite previous patents that are in closely related technological fields to her own, and one might expect that inventors working in the same field are more likely to live in the same country. Again, we would like to be able to measure the geographic localization effect while controlling for technological localization effects. The econometric model that we develop below is meant to allow us to sort out and measure each of these different effects.

III. THE DATA

We are in the final stages of collecting from commercial sources a complete database on all U.S. patents[1] granted since 1963. It includes data for each patent indicating the nature of the organization, if any, to which the patent property right was assigned; the names of the inventors and the organization, if any, to which the patent right was assigned; the residence of each inventor[2]; the date of the patent application and the patent grant; a detailed technological classification for the patent; and miscellaneous other information. A file indicating all of the citations made by U.S. patents since 1977 to previous U.S. patents complements the data on individual patents. Using the citation information in conjunction with the detailed information about each patent itself, we have a rich mine of information about individual inventive acts and the links among them as indicated by citations made by a given patent to a previous one.

We have discussed elsewhere at great length the advantages and disadvantages of using patents and patent citations to indicate inventions and knowledge links among inventions (Jaffe, Henderson and Trajtenberg, 1993; Trajtenberg, Henderson and Jaffe, 1997; see

[1] By "U.S. patents," we mean in this context patents granted by the U.S. patent office. All of our research relies on U.S. patents in this sense. Currently about half of U.S. patents are granted to foreigners.

[2] City and state for U.S. inventors, country for non-U.S. inventors.

also Griliches, 1990). Patent citations perform the legal function of delimiting the patent right by identifying previous patents whose technological scope is explicitly placed outside the bounds of the citing patent. Hence the appearance of a citation indicates that the cited patent is, in some sense, a technological antecedent of the citing patent. Patent applicants bear a legal obligation to disclose any knowledge that they might have of relevant prior inventions, and the patent examiner may also add citations not identified by the applicant. In related work, Jaffe, Fogarty and Banks (forthcoming) examined in detail the patents that did and did not contain citations to a specific set of important NASA patents. The conclusion was that, while citations contain much "noise" in the form of apparently spurious implied connections, on the whole they do provide useful information about the generation of future technological impact of a given invention.

The analysis in this paper is based on citations to patents granted between 1963 and 1993. We examine a set of "potentially cited" patents whose primary inventor resided in the U.S., Great Britain, France, Germany or Japan, and which were assigned to corporations. There are a total of about 1.5 million such patents. About 65% of these are from the U.S. (i.e.., their inventors reside in the U.S.), 17% are from Japan, 10% from Germany, 5% from Great Britain, and 4% from France. We then examine all citations made to these corporate patents (whether or not the citing patent is itself corporate) from patents granted in any of these five countries between 1977 and 1994. There are about 1.2 million "citing" patents, and they made a total of about 5.0 million citations to the set of "potentially cited" patents that we are considering.

Patenting in different countries differs in ways that affect observed citation frequencies. Some indications of these differences are presented in Table 1, which summarizes the patent data, from both the cited and citing perspectives, for the five different countries. We also classify patents into five broad technological fields, based on the main patent class assigned to the patent by the patent examiner.[3] These fields are: Drugs and Medical Technology; Chemicals, excluding Drugs; Electronics, Optical and Nuclear Technology;

Mechanical Technology; and All Other. Table 1 gives some totals for patents and citations for each of the five countries and five technology fields. Overall, 6% of the cited patents are in Drugs and Medical, 28% in Chemicals, 22% in Electronics, etc., 35% in Mechanical and 9% All Other. There are, however, significant variations across the countries in the field composition of their patents. In particular, Japan has a larger share of electronic patents and Germany a larger share of chemical patents than the U.S. Since citation intensities vary by field within countries, raw differences between countries as in Figure 1 are a mixture of country effects, field effects, and field-country interaction effects. We discuss below how to sort out these effects.

Table 1
Patents and Citations by Country

	Potentially Cited Patents	Average Citations per Patent	Fraction of Self-Cites	Potentially Citing Patents
United States				
1969	36406	2.74	0.07	
1993	36512	0.31	0.31	53235
average per year	31135	3.50	0.29	42494
Drugs and Medical	5%	4.54	0.22	
Chemical, exc. Drugs	29%	3.35	0.34	
Electronics, Optics and Nuclear	21%	3.84	0.26	
Mechanical	34%	3.32	0.30	
All Other	11%	3.30	0.25	
Great Britain				
1969	2713	1.98	0.02	
1993	1995	0.23	0.16	2294
average per year	2223	2.86	0.22	2471
Drugs and Medical	8%	3.10	0.22	
Chemical, exc. Drugs	29%	2.89	0.24	
Electronics, Optics and Nuclear	20%	3.17	0.17	
Mechanical	35%	2.66	0.23	
All Other	9%	2.64	0.18	

[3] There are about 400 of these patent classes.

	Potentially Cited Patents	Average Citations per Patent	Fraction of Self-Cites	Potentially Citing Patents
France				
1969	1341	2.04	0.01	
1993	2446	0.22	0.24	2908
average per year	1673	2.58	0.18	2457
Drugs and Medical	8%	2.40	0.19	
Chemical, exc. Drugs	28%	2.42	0.20	
Electronics, Optics and Nuclear	22%	2.84	0.14	
Mechanical	35%	2.63	0.18	
All Other	8%	2.40	0.16	
Germany				
1969	3785	2.15	0.06	
1993	6255	0.21	0.27	6891
average per year	4894	2.64	0.22	6603
Drugs and Medical	6%	2.72	0.20	
Chemical, exc. Drugs	34%	2.56	0.27	
Electronics, Optics and Nuclear	17%	2.73	0.18	
Mechanical	35%	2.71	0.20	
All Other	8%	2.42	0.19	
Japan				
1969	1758	2.02	0.05	
1993	20997	0.31	0.28	22291
average per year	8040	3.44	0.19	13780
Drugs and Medical	5%	2.80	0.18	
Chemical, exc. Drugs	23%	3.21	0.22	
Electronics, Optics and Nuclear	30%	3.72	0.19	
Mechanical	36%	3.54	0.18	
All Other	6%	2.82	0.19	
All				
1969	46003	2.60	0.04	
1993	68205	0.30	0.25	87619
average per year	47965	3.34	0.22	67805
Drugs and Medical	6%	3.89	0.20	
Chemical, exc. Drugs	28%	3.18	0.25	
Electronics, Optics and Nuclear	22%	3.66	0.19	
Mechanical	35%	3.24	0.22	
All Other	9%	3.12	0.19	

Table 1 also shows that a significant fraction of citations from each country are "self-citations." Self-citations are defined as those for which the citing and cited patent are both assigned to the same corporate organization.[4] Self-citations are more common in the U.S. than in other countries. It also turns out that self-citation come more quickly on average, and are more geographically localized. In order to get measures that more closely correspond to knowledge "spillovers," most of the analysis below is carried out excluding these self-citations.

Finally, the number of patents taken out in the U.S. has grown at dramatically different rates for different countries. In particular, while the number of U.S. invented patents in 1993 was essentially equal to the number in 1969, the number of Japanese-invented patents increased by 1194% over that same period, and the number of Great Britain-invented patents declined by about 26%.[5] Thus when we compare overall citation frequencies for the different countries, we are looking at averages which are tilted towards different citing cohorts.

IV. MODELING

A. Patent-pair citation frequencies

We seek to model the citation frequencies described in Section II above, the way in which these frequencies evolve over time, and how they are affected by characteristics of the citing and cited patent. One way to approach this would be with a probit-type model, in which each citation is an observation, and the regression dataset is created by combining the actual citations with a random sample of patent pairs that did not cite each other. One could then ask how the predicted probability that a patent pair will result in a citation is affected by various regressor variables.[6]

[4] Identifying self-citations is complicated by the fact that patents may be assigned to corporate entities that are affiliates or subsidiaries of other entities. We are in the process of refining the corporate assignments of the patents in the database to take this into account, but the results in this paper are based on self-citation defined only in terms of patents assigned to the same corporate entity. This almost certainly understates the extent of self-citation.

[5] As a reminder, these numbers are for patents taken out in the U.S. and assigned to a corporation; U.S.-invented versus Japanese-invented is determined on the basis of the postal address of the primary inventor.

[6] See Podolny and Shepard, 1997, for an application of this approach.

In this application, however, we observe approximately five million citations; if this were combined with an equal number of non-citing patent pairs, the regression dataset would have ten million observations. The number of unique combinations of values of potential regressor variables is, however, a small fraction of that. Put differently, if one were to run a probit with those ten million observations, very many of those observations would have identical values for any conceivable set of right-hand-side variables. In such a case, no information is lost by combining observations into "cells" characterized by the values of the regressor variables, and making the dependent variable the *fraction* of the patent pairs in the cell for which a citation occurred. In this way, we reduce the number of observations from more than five million (the exact value depending on the sampling from the non-citing pairs) into a dataset with about 50,000 observations, with little loss of relevant information.

Most of our potential regressors are categorical rather than continuous variables, such as cited country, citing country, technology field, cited year and citing year. In addition to these effects, we wish to capture the evolution of citations over elapsed time as shown in Figure 1. For this purpose we adapt the formulation of Caballero and Jaffe (1993) and Jaffe and Trajtenberg (1996). The citation frequency (the likelihood that any particular patent K granted in year T will cite some particular patent k granted in year t) is assumed to be determined by the combination of an exponential process by which knowledge diffuses and a second exponential process by which knowledge becomes obsolete. That is:

$$p(k,K) = [1 + \gamma D(k,K)]\alpha(k,K)\exp[-\beta_1(k,K)(T - t)][1 - \exp(-\beta_2(T - t))] \qquad 1$$

where β_1 determines the rate of obsolescence and β_2 determines the rate of diffusion. The parameter α is a shift parameter that depends on the attributes of both the patent *k* and the patent *K*. D(k,K) is a dummy variable, set equal to unity if the patent k is in the same patent class as the patent K, and zero otherwise. Thus, the parameter γ measures the overall increase in citation frequency associated with the two patents matching by patent class. The dependence of the parameters α and β_1 on k and K is meant to indicate that these could be functions of certain attributes of both the cited and citing patents. In this paper, we consider the following as attributes of the cited patent k that might affect its citation frequency:

- t, the grant year of the potentially cited patent,
- ℓ, the "location" of the cited inventor (U.S., Great Britain, France, Germany or Japan),
- g=1...5, the technological field of the potentially cited patent.

As attributes of the potentially citing patent K that might affect the citation likelihood we consider:

- T, the grant year of the potentially citing patent, and
- L=1...5, the location of the potentially citing patent.

Additional insight into this parameterization of the diffusion process can be gained by computing the lag at which the citation function is maximized ("the modal lag"), and the modal probability of citation. A little calculus shows that the modal lag is approximately equal to $1/\beta_1$; increases in β_1 shift the citation function to the left. The maximum value of the citation frequency is approximately determined by β_2/β_1. Increases in β_2 holding β_1 constant increase the overall citation intensity,[7] and are roughly equivalent to increasing the citation frequency *proportionately* at every value of (T-t). That is, variations in β_2, holding β_1 constant are not separately identified from variations in α. Thus, since the model is somewhat easier to estimate and interpret with variations in α, we do not allow variations in β_2.

B. Expected citation count for "cells"

Consider a potentially cited patent with particular t, ℓ, g attributes, e.g., a Japanese-invented patent in the Drug and Medical area granted in 1985. The expected number of citations that this patent will receive from a particular patent with a given T,L combination (e.g., a British patent granted in 1993 that happens to be in the same patent

[7] The approximation involved is that $\log(1 + \beta_2/\beta_1) \approx \beta_2/\beta_1$. Our estimations all lead to β_2/β_1 on the order of 10^{-6}, and indeed the approximation holds to five significant figures.

fields, country-year pairs will vary in the extent of overlap in patent classes class) is just the above likelihood, as a function of ℓ,t,g,T,L and D(k,K). The expected number of citations *from all* patents with a given T,L combination is found by summing the frequency shown in Eq. **1** over all such patents. Similarly, the expected total number of citations *to all* patents with the particular ℓ,t,g combination will be found by summing over all such patents. The only tricky part of this double summation is dealing with D(k,K). We show in Appendix A that one can start from Eq. (1) and aggregate to derive a relationship for "cells" identified by ℓ,t,g,T and L, where the dependent variable is the expected frequency of citation $p_{ltgTL} \equiv \dfrac{C_{ltgTL}}{(N_{TL})(N_{ltg})}$, i.e., the ratio of the number of citations to the product of the number of potentially citing and potentially cited patents. In expectation, this frequency is a function of the characteristics of k and K, and the variable:

$$PROX_{\ell tgTL} = \sum_s f_{\ell tgs} f_{TLS}$$

where $f_{\ell gs}$ is the fraction of potentially cited patents in patent class s and f_{TLS} is the fraction of potentially citing patents in patent class s. PROX measures the extent to which the potentially citing and potentially cited patents overlap in their patent class distribution.[8] It is closely related to the technological proximity measure of firms used in Jaffe (1986). This brings us to the following equation:

$$p_{ltgTL} = \alpha_{ltgTL}\left[1+\gamma PROX_{ltgTL}\right]\exp[-(\beta_1)_{ltgTL}(T-t)][1-\exp(-\beta_2(T-t))] + \varepsilon_{lgtTL} \quad \textbf{(2)}$$

which can be estimated by non-linear least squares if the error $\varepsilon_{\ell gtTL}$ is well-behaved.

The data set consists of one observation for each feasible combination of values of ℓ,t,g,T and L. Since t runs from 1963 to 1993 and T runs from (the greater of 1977 and t+1) to 1994, the number of cells for each ℓ, g, L combination is 14x18 + (17+16+15+14+13...+1). There are 125 ℓ, g, L combinations, so the total number of cells is 50,625. Simple statistics for this dataset are presented in Table 2. The average number of cited patents in a cell is about 1800; the minimum is 16 (French Drug and Medical

[8] Recall that there are about 400 such patent classes, so that even within the five broad technology fields, country-year pairs will vary in the extent of overlap in patent classes.

patents in one particular year) and the maximum is almost 15,000 (U.S. Mechanical patents in one particular year). The number of citations varies from 0 to over 6000 with a mean of about 100; the mean of the citation frequency is about 4×10^{-6}.

Table 2
Statistics for Regression Variables

Label	Mean	Std Dev	Minimum	Maximum
number of citations	98	341	0	6326
number of non-self-cite citations	78	270	0	5010
potentially cited patents	1791	2928	16	14735
potentially citing patents	14164	16141	1610	56065
cited grant year	1974.4	7.4	1963	1993
citing grant year	1986.7	5.0	1977	1994
citation frequency ($\times 10^6$)	3.85	5.34	0	123.24
non-self-cite citation frequency ($\times 10^6$)	2.78	3.11	0	40.57
technological proximity of cells	0.0075	0.0069	0.0008	0.0620
lag in years	12.3	7.4	1	31
regression weight (square root[ncited*nciting])	3391	3745	160	28742

C. Econometric issues and interpretation

The first specification issue to consider is the difficulty of estimating effects associated with cited year, citing year and lag. This is analogous to estimating "vintage," time, and age effects in a wage or a hedonic price model. If lag (our "age" effect) entered the model linearly, then it would be impossible to estimate all three effects. Given that lag enters our model non-linearly, all three effects are, in principle, identified. In practice, however, we found that we could not get the model to converge with the double-exponential lag function and separate α parameters for each cited year and each citing year. We were, however, able to estimate a model in which cited years are grouped into five-year periods, indexed by p. Hence we assume that $\alpha(t)$ is constant over t within these periods, but allow the periods to differ from each other.

The estimation is carried out including "base" values for β_1 and β_2. Location, field, cited period and citing year effects are all estimated relative to a base value of unity.[9] The various different effects are included by entering multiplicative parameters, so that the estimating equation looks as follows:

$$
\begin{aligned}
p_{ltgTL} = {} & \left[1 + \gamma PROX_{ltgTL}\right] \alpha_t \alpha_g \alpha_T \alpha_{lL} \\
& \exp[-(\beta_1)\beta_{1g}\beta_{1lL}(T-t)][1-\exp(-\beta_2(T-t))] \; + \; \varepsilon_{lgtTL}
\end{aligned} \tag{3}
$$

Thus we allow α to vary by cited period, cited field, citing year and all possible combinations of citing country and cited country. We allow β_1 to vary by cited field, and every possible combination of cited and citing countries. In this model, unlike the linear case, the null hypothesis of no effect corresponds to parameter values of unity rather than zero (except for γ, and the "base" values of β_1 and β_2). For each effect, one group is omitted from estimation, i.e., its multiplicative parameter is constrained to unity. Thus the parameter values are interpreted as relative to that base group.[10]

The estimate of any particular $\alpha(k)$, say α(g=Chemical), is a proportionality factor measuring the extent to which the patents in the Chemical field are more or less likely to be cited over time vis à vis patents in the base category (Drugs). Thus, an estimate of α(g=Chemical) = 1.5 means that the likelihood that a patent in the field of Chemicals will receive a citation is 50% higher than the likelihood of a patent in the base category, controlling for other factors. Notice that this is true across all lags; we can think of an α greater than unity as meaning that the citation function is shifted upward proportionately, relative to the base group. Hence the integral over time (i.e., the total number of citations per patent) will also be 50% larger. Similarly, if α (ℓ = Japan, L=U.S.) is .72, this means that a Japanese patent is 28% less likely to get a citation from a random U.S. patent than is a random U.S. patent.

---

[9] As noted above, α is not separately identified from β_1 and β_2. Hence we do not estimate a "base" value for the parameter α; it is implicitly unity.

[10] The base group for each effect is: Cited time period (p), 1963-65; Cited field (g), "Drugs and Medical"; Citing year (T), 1977; and Cited/Citing country (ℓ L),U.S.-citing-U.S.

We can think of the overall citation intensity measured by variations in α as composed of two parts. Citation intensity is the product of the "fertility" (Caballero and Jaffe, 1993) or "importance" (Trajtenberg, Henderson and Jaffe, 1997) of the underlying ideas in spawning future technological developments, and the average "size" of a patent, i.e., how much of the unobservable advance of knowledge is packaged in a typical patent. Within the formulation of this paper, however, it is not possible to decompose the α-effects into these two components.[11]

In the case of $\alpha(K)$, that is, when the multiplicative factor varies with attributes of the citing patents, variations in it should be interpreted as differences in the probability of *making* a citation, all else equal, for patents in a particular category vis à vis the base category. If, for example, $\alpha(\ell = \text{U.S.}, L = \text{Japan})$ is 0.76, this means that the average patent granted to Japanese inventors is three-quarters as likely as a patent granted to inventors residing in the U.S. to cite any given U.S. patent. Note that, just as variations in α across cited patents are composed of both variations in fertility or importance and variations in "patent size," variations across citing patents can be caused by both variations in true "knowledge use" and variations in the "propensity to cite." Because there are institutional reasons why the propensity to cite may vary across countries, this has important consequences for interpreting the results. We return to this issue below.

Variations in β_1 (by attributes of either the cited or the citing patents) imply differences in the timing of citations across categories of patents. Higher values of β_1 mean higher rates of decay, which pull the citations function downwards and leftward. In other words, the likelihood of citations would be lower everywhere for higher β_1, and would peak earlier on. Thus a higher α means more citations at all lags; a lower β_1 means more citations at later lags.

[11] Caballero and Jaffe (1993) attempt to identify the size of patents by allowing exponential obsolescence to be a function of accumulated patents rather than elapsed calendar time. We intend to explore this possibility in future work.

When both α and β_1 vary, the citation function can shift upward at some lags while shifting downward at others. For example, if α for citations from Japan to Japan is 2.32 and the β_1 for Japan to Japan is 1.54, this implies that the likelihood of citation in early years is higher than the base group, but because of the higher β_1, this difference fades over time. Because obsolescence is compounded over time, differences in β_1 eventually result in large differences in the citation frequency.[12] If we compute the ratio of the likelihood of citations for Japan-to-Japan relative to U.S.-to-U.S. using these parameters, we find that one year after being granted, Japan-to-Japan citations are about twice as likely as U.S.-to-U.S., but nine years down the road the frequencies for the two groups are about the same, and at a lag of 20 years Japan-to-Japan citations are actually about 70% *less* likely than for the base category.

A final interpretation issue relates to citations from patents assigned to the same firm as is the cited patent, so-called "self-citations." As discussed by Jaffe, Henderson and Trajtenberg (1993), self-citations cannot be regarded as evidence of spillovers. Hence if we are interested in geographic localization of spillovers, we want to exclude self-citations. On the other hand, self-citations are an important indicator of the cumulative nature of technology, and of firms' ability to appropriate the returns to their inventions.[13] Thus for some purposes, such as assessing the overall role of technology in regional economic development, localization of citations inclusive of self-citations is of interest. In order to focus on spillovers, we concentrate on the results exclusive of self-citations, but we comment briefly on the very high degree of localization of self-citations.

[12] Since increases in β1 both reduce the peak frequency (β_2/β_1) and cause the function to decay from the peak faster, such increases reduce the cumulative citation frequency or integral over time non-linearly. The cumulative frequency is approximately $\alpha\ \beta_2/(\beta_1)^2$.

[13] The extent of self-citation is an indicator of firms' successful appropriation. See Trajtenberg, Jaffe, and Henderson, 1997. Putnam (1997) finds that the number of self-citations is a good predictor of firms' decision to pay renewal fees for patents that would otherwise expire.

We estimate Eq. **3** by non-linear least squares. Since the left-hand variable is an empirical frequency, the model is heteroskedastic. To improve efficiency and get the right standard errors, we weight the observations by the reciprocal of the estimated variance, $\sqrt{(N_{ltg})(N_{LT})}$. In general, this weighting greatly improves the fit of the model, but does not alter the parameter estimates materially.

IV. RESULTS

Complete results from the estimation of Eq. **3** are presented in Appendix B. The model has 82 parameters (γ, base values of β_1 and β_2; 24 cited country/citing country interactions for α; 4 technology field effects for α; 6 cited time period effects for α; 17 citing year effects for α; 24 cited country/citing country effects for β_1; 4 technology field effects for β_1). Overall, the model fits the data reasonably well. Because of the large sample size, the estimated standard errors are quite small. The base value for β_1 is about .2, suggesting a modal lag of about five years, which is not surprising based on Figure 1. The estimate for the technology match parameter γ is 99, which means that a patent is about 100 times more likely to cite a patent in the same patent class as it is to cite a random patent in some other class. In reality, of course, some classes are "closer" to each other than others in technology space, but it is not surprising that, on average, patents in the same class are much more likely to cite each other than to cite patents in any of the other classes.[14]

Technology field effects are present in both the α's and the β_1's, but the β_1 effects are not large. The α's greater than 1 mean that all other fields receive more citations than Drug and Medical patents (the base group). The β_1's greater than 1 mean that other fields receive citations somewhat faster than Drugs.[15] The cited time period and citing year

---

[14] Podolny and Shepard (1997), looking only at patents within a group of classes related to semiconductor technology, found that citations from patents in the same class where about 15 times as likely.

[15] The field differences found here for β_1 are smaller than we found in our previous paper (Jaffe and Trajtenberg, 1996). Since that paper looked only at citations to U.S. patents, this suggests that there is a greater variation in citation speed across fields in U.S. patents than in other countries' patents. In particular, we found that the citations to U.S. patents came much faster in electronics than in the other fields; this effect is present in the overall data but is not as big.

effects are similar to what we have found before: the number of citations received rises in the 1970s and 1980s, and the number of citations made rises essentially throughout the whole period.

Table 3 presents the estimates of the α and β_1 parameters in several different ways. The top panel simply reproduces the α estimates presented in Appendix B, but arrays them in matrix form. The second panel presents the estimated values (with standard errors) in terms of $(1/\beta_1)$, which has years as units and is equal to the lag at which the citation frequency reaches its maximum value. The bottom panel presents estimated values (with standard errors) for $\alpha\beta_2/(\beta_1)^2$, which is the integral of the citation function from t=0 to infinity. This is an estimate of the expected number of citations that a single patent will receive from a set of patents consisting of one random patent per year forever. Thus the middle panel of the table measures the "speed" of citation diffusion and the bottom panel measures the overall intensity of citation.[16]

Table 3
Regression Coefficients in Matrix Form

Alphas

	Citing				
	United States	Great Britain	France	Germany	Japan
Cited					
United States	1.00	0.72	0.65	0.56	0.76
Great Britain	0.71	1.78	0.79	0.75	0.66
France	0.60	0.72	2.17	0.73	0.63
Germany	0.55	0.73	0.74	1.32	0.83
Japan	0.72	0.62	0.67	0.81	2.32

[16] We also tested various restricted versions of this model to see if the parameter differences reported here are jointly significant. The following restricted versions of the model were all rejected, with p-values of .0001 or less in the appropriate chi-squared test, in favor of the reported model: country effects in the α's but not the β_1's; effects for each cited and citing country, plus a "domestic localization" effect common to all countries, but no cited country/citing country interaction effects; cited country/citing country interaction effects in the α's but only cited and citing country effects plus the localization effect in the β_1's.

Modal Lag

	Citing				
	United States	Great Britain	France	Germany	Japan
Cited					
United States	5.25	5.08	5.08	5.08	4.41
	(0.049)	(0.070)	(0.077)	(0.074)	(0.051)
Great Britain	5.39	4.24	4.63	4.68	4.46
	(0.073)	(0.049)	(0.103)	(0.092)	(0.088)
France	5.43	4.85	4.02	4.67	4.51
	(0.087)	(0.114)	(0.047)	(0.101)	(0.098)
Germany	5.42	4.56	4.58	4.23	4.14
	(0.079)	(0.093)	(0.096)	(0.051)	(0.062)
Japan	4.99	4.80	4.52	4.45	3.40
	(0.061)	(0.104)	(0.102)	(0.070)	(0.030)

Cumulative Probability

	Citing				
	United States	Great Britain	France	Germany	Japan
Cited					
United States	1.49	1.01	0.91	0.78	0.80
	(0.109)	(0.075)	(0.068)	(0.058)	(0.059)
Great Britain	1.11	1.72	0.92	0.88	0.71
	(0.082)	(0.128)	(0.070)	(0.066)	(0.053)
France	0.95	0.91	1.89	0.85	0.70
	(0.071)	(0.070)	(0.141)	(0.065)	(0.053)
Germany	0.86	0.82	0.84	1.28	0.76
	(0.064)	(0.062)	(0.064)	(0.095)	(0.057)
Japan	0.97	0.77	0.74	0.87	1.45
	(0.071)	(0.059)	(0.056)	(0.065)	(0.108)

Several features of these matrices are worth noting. Looking first at the α's, the diagonal elements strongly dominate both the rows and columns of the matrix. What this means is that there is a strong pattern of geographic localization, in the sense that the domestic citation function is shifted upward. This is true for all countries, and it is true whether one compares the domestic citations to citations *received from* other countries (across the rows) or citations *made to* other countries (down the columns). The other notable feature of the top panel of Table 3 is the symmetry of the matrix. For example, α for Germany citing U.S. and for U.S. citing Germany are the two lowest numbers in the matrix. Conversely, the two highest non-diagonal numbers in the α table are for Germany citing Japan and Japan citing Germany. Although these differences among the off-diagonal elements are not as large as the localization effect of domestic citation, it suggests that inter-country knowledge flows are typically bi-directional, with relatively large or small flows in one direction being associated with similar flows in the other direction.

Geographic localization is also evident in the β_1 parameters, presented in the middle panel of Table 3 in the form of the estimated modal lag. Here the diagonal elements are generally the *smallest* entry in each row and column, meaning modal citation lags are noticeably shorter for domestic citations, relative to citations to and from others. The only exception to this general pattern is the U.S. U.S. inventors are slightly faster to cite Japanese inventors than they are to cite U.S. inventors (β_1=1.05), and Japanese inventors are faster to cite U.S. inventors than are U.S. inventors (β_1=1.19).

There are also systematic variations across the countries that are superimposed on top of the general pattern of localization. While the modal lags for citations made by the U.S. range from 5 to 5.4 years (depending on the cited country), those for Japan range from 3.4 to 4.4. Indeed, it appears that the overall tendency of the U.S. to both generate and receive long-lagged citations is part of the reason why U.S.-to-U.S. citations do not come more quickly than those from and to others.

The fact that domestic citations generally involve both higher α and higher β_1 creates offsetting effects for the overall number of citations, since the higher β_1 means that citations fade faster and hence reduces the total holding α constant. The bottom panel of Table 3 combines these effects by presenting the overall cumulative probabilities. The estimates show that, in terms of total citations, the variations in α dominate the variations in β_1; the matrix is still strongly diagonal, indicating localization. These differences are quite significant statistically, although it should be noted that this calculation relies heavily on the assumed functional form as it integrates the citation function into the infinite future.

Also noticeable in the cumulative probabilities is that the U.S. tends to both make and receive more citations than other countries. Note, for example that the entry in the U.S.

column contains the largest figure other than the diagonal in every row, and the U.S. row contains the largest figure other than the diagonal in every column except Germany. This result could be driven by differences between the U.S. and other countries in the propensity to patent. If the U.S. has a low propensity to patent, then each patent granted represents (on average) a larger chunk of knowledge, which could result in more citations made and received per patent (Cabellero and Jaffe, 1993). It is more likely, however, that the propensity of U.S. inventors to patent *in the U.S.* is greater than that of foreigners (Eaton and Kortum, 1996). That is, U.S. inventors are more likely to take out a U.S. patent on a trivial invention than are foreigners. All else equal, this should make the average citation rate to and from the U.S.-invented patents *lower* than the corresponding rates for foreign-invented patents. Since we find the opposite, this may be evidence confirming a view of the U.S. as the most open and interconnected economic and technological system.

There are also some interesting pairwise effects. The U.S. and Great Britain are "closer" to each other (in terms of overall probability of citation) than any other country pair, suggesting a possible effect of common language. Japan is closer to the U.S. than it is to any of the European countries. Britain and France are closer to each other than to Germany, but closer to Germany than to Japan. Note, however, that not all of these differences are statistically significant.

Figures 2 through 6 show the effects of these parameter differences graphically, and also present a useful pictorial comparison to the raw data presented in Figure 1. Each Figure presents the estimated citation functions for citations *to* one of the countries, with the different lines within each Figure corresponding to the different citing countries. Comparing Figure 2 to Figure 1 shows some of the effects of controlling for non-geographic effects. First, as suggested above, the "tails" in the estimated functions in Figure 2 are much thinner. Second, while geographic localization is clearly present in Figure 2, its magnitude is noticeably diminished, with the citation frequency for other countries at the modal lag being roughly 55-75% of U.S.-U.S. as compared to 40-60% in

Figure 1. As discussed further below, this is primarily the effect of eliminating self-citations.

Figure 2
Estimated Citation Functions for Citation to U.S.-Invented Patents (Excluding Self-citations)

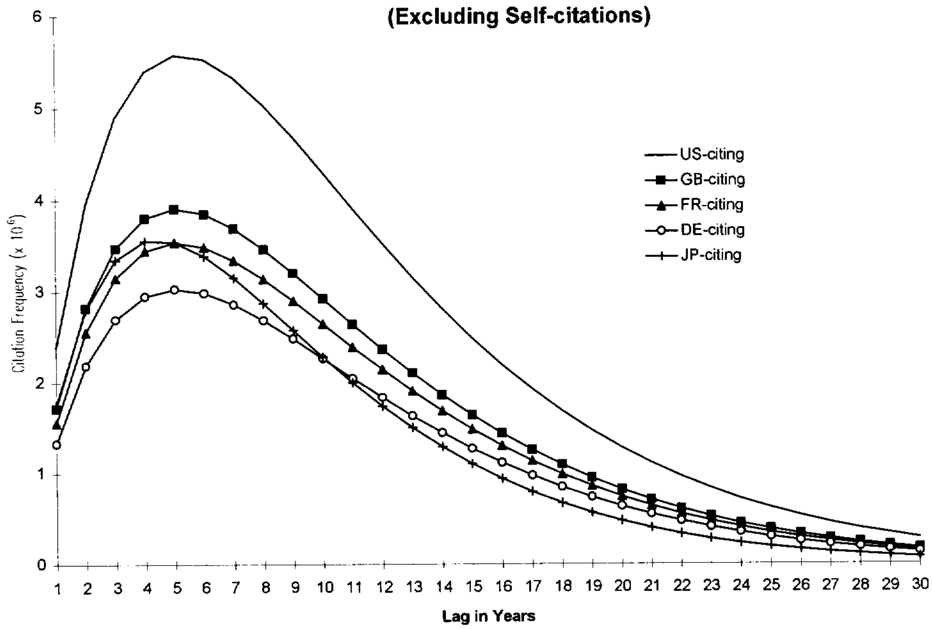


Figure 3
Estimated Citation Functions for Citation to British-Invented Patents (Excluding Self-citations)

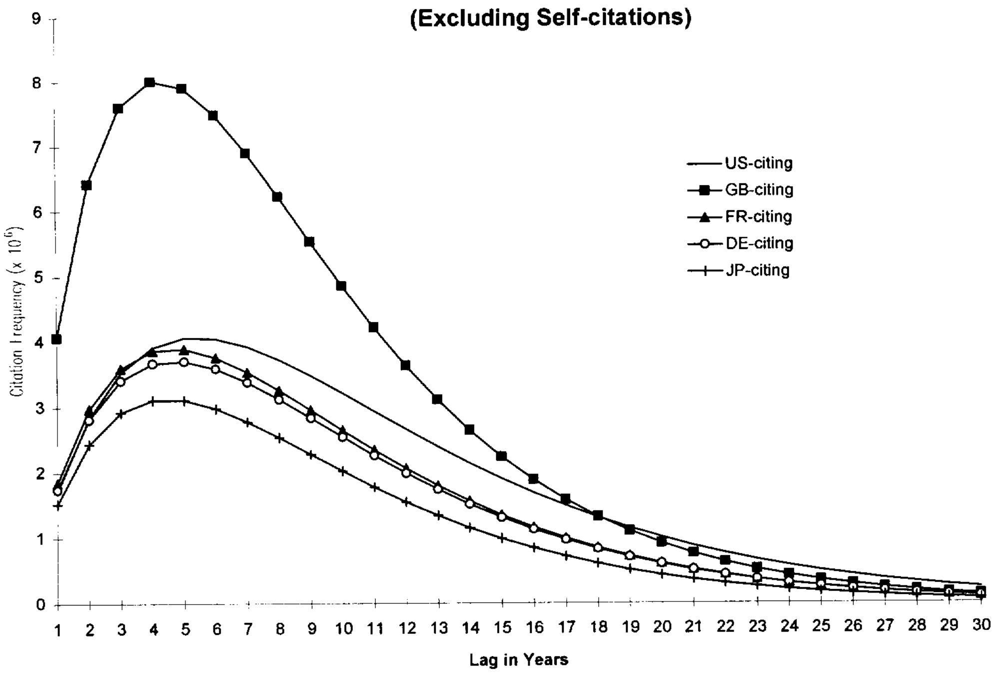


Figure 4
Estimated Citation Functions for Citation to French-Invented Patents (Excluding Self-citations)

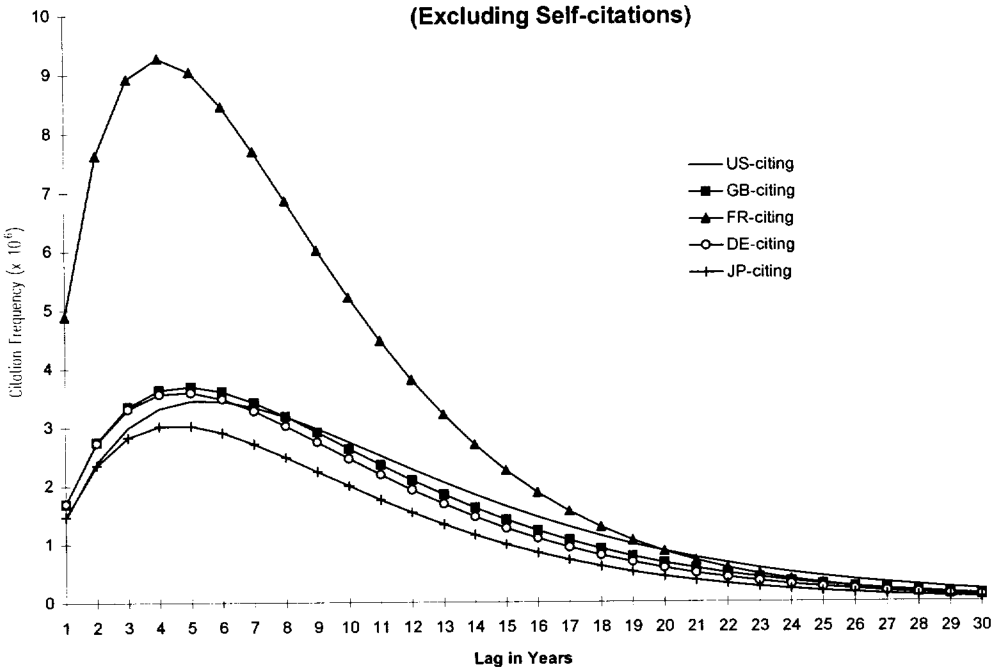


Figure 5
Estimated Citation Functions for Citation to German-Invented Patents (Excluding Self-citations)

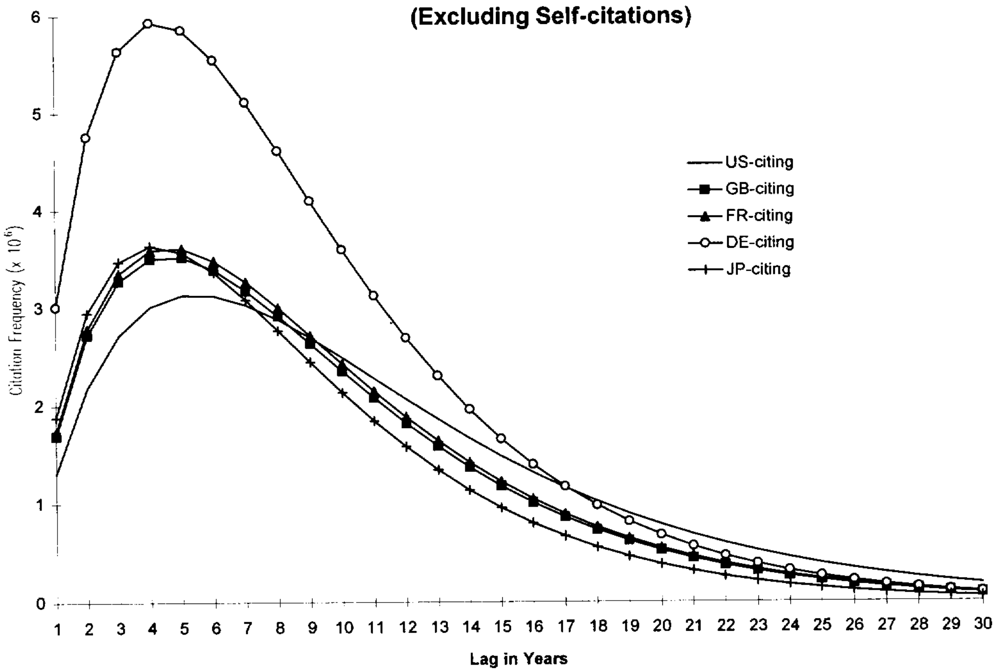
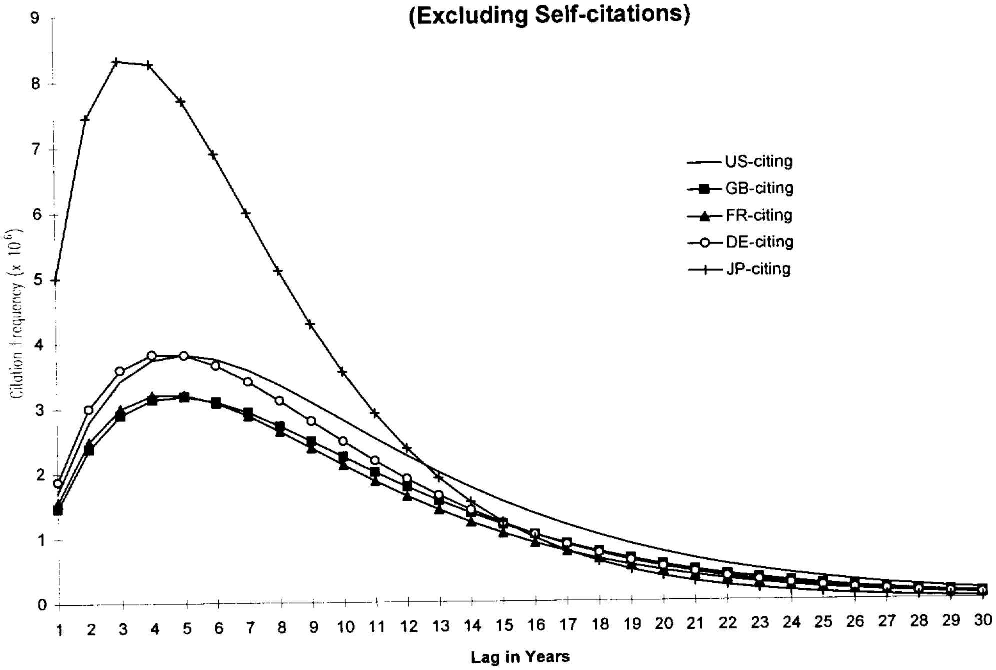


Figure 6
Estimated Citation Functions for Citation to Japanese-Invented Patents
(Excluding Self-citations)



In terms of the effects seen numerically in Table 3, the Figures show the "speed" of Japan, as its line typically peaks early and then fades, and the "slowness" of the U.S., whose predicted frequency of citation is the highest after long lags in all of the pictures. The graphs also show that the differences among non-domestic citing countries are always smaller than the localization effect that separates domestic citations from foreign ones.

Figures 2 to 6 generally show a pattern of "fading" of geographical localization. The combination of relatively high α and relatively high β_1 for domestic citations means that the initial domestic probability is much higher, but that it fades faster, so that other countries typically catch up eventually. This can be seen in the "crossing" of the domestic citation function with the others after 15 to 25 years. This phenomenon is also illustrated in Table 4, which gives the probability of citation from various countries *relative to* the domestic citation probability, for each cited country, in the first year and after 20 years.

For every cited country except the U.S., the relative citation frequency of the other countries is greater after 20 years than in the first year. Indeed, for Japan, every other country cites its twenty-year-old patents with greater frequency than it does itself. This results from the combined effect of fading of localization and the fact that Japan is generally a high-β_1 (fast-fading) maker of citations. Conversely, the lack of fading of geographic localization in citations to the U.S. reflects the general tendency of the U.S. toward low β_1 (slow fading).

Table 4
Fading of Geographic Localization over Time

Relative Citation Rate for Citations to U.S.:

	Citing Country				
Lag in years	United States	Great Britain	France	Germany	Japan
1	1.00	0.72	0.65	0.56	0.73
20	1.00	0.64	0.58	0.49	0.37

Relative Citation Rate for Citations to Great Britain:

	Citing Country				
Lag in years	United States	Great Britain	France	Germany	Japan
1	0.42	1.00	0.45	0.43	0.37
20	1.10	1.00	0.67	0.65	0.47

Relative Citation Rate for Citations to France:

	Citing Country				
Lag in years	United States	Great Britain	France	Germany	Japan
1	0.29	0.34	1.00	0.35	0.30
20	1.00	0.77	1.00	0.67	0.50

Relative Citation Rate for Citations to Germany:

	Citing Country				
Lag in years	United States	Great Britain	France	Germany	Japan
1	0.44	0.56	0.57	1.00	0.62
20	1.16	0.77	0.81	1.00	0.56

Relative Citation Rate for Citations to Japan:

	Citing Country				
Lag in years	United States	Great Britain	France	Germany	Japan
1	0.34	0.29	0.31	0.38	1.00
20	2.01	1.49	1.23	1.39	1.00

All of the results discussed so far derive from estimation of the "full" model of Eq. **3**. For comparison to our earlier paper, as well as for the light it sheds on the interaction of different effects, it is useful to consider briefly how the results differ in less complete or different models. In particular, our earlier research did not exclude self-citations, and did not include the "technological proximity" effect. These effects are interesting in their own right, and may also be expected to interact in important ways with geographic localization. Table 5 summarizes the results with and without these non-geographic effects. Generally, excluding self-cites significantly reduces the apparent geographic localization, as well as reducing the extent to which that localization "fades." That is, the citation intensity from other countries, relative to the domestic citation rate, is lower in columns 1 and 2 than in columns 3 and 4 in the first year, but is higher in columns 1 and 2 than in columns 3 and 4 after 20 years. What this means is that self-cites are highly geographically localized (which should not be a surprise) and generally come at shorter lags (Trajtenberg, Henderson and Jaffe, 1997). Thus including them creates strong localization particularly in early years; excluding them dilutes localization; this weaker initial localization then also fades less.

Table 5
Comparison of Models

	No Tech. Proximity Parameter, Incl. Self-Cites	With Tech. Proximity Parameter, Incl. Self-Cites	No Tech. Proximity Parameter, Excl. Self-Cites	With Tech. Proximity Parameter, Excl. Self-Cites
R^2	0.779	0.813	0.746	0.765
Beta1	0.205	0.208	0.191	0.190
Beta2	5.258×10^{-6}	2.019×10^{-6}	7.309×10^{-7}	2.891×10^{-7}
Technological Proximity Parameter	n.a.	101.24	n.a.	99.49
Citations to U.S.				
Citation Intensity Relative to U.S.-U.S.				
Year 1				
Great Britain	0.54	0.50	0.75	0.72
France	0.45	0.44	0.66	0.65
Germany	0.36	0.37	0.54	0.56
Japan	0.49	0.51	0.72	0.73
Year 20				
Great Britain	0.80	0.78	0.67	0.64
France	0.73	0.74	0.59	0.58
Germany	0.65	0.67	0.50	0.49
Japan	0.44	0.44	0.38	0.37

	No Tech. Proximity Parameter, Incl. Self-Cites	With Tech. Proximity Parameter, Incl. Self-Cites	No Tech. Proximity Parameter, Excl. Self-Cites	With Tech. Proximity Parameter, Excl. Self-Cites
Citations to France Citation Intensity Relative to France-France				
Year 1				
U.S.	0.09	0.11	0.26	0.29
Great Britain	0.14	0.12	0.36	0.34
Germany	0.11	0.12	0.31	0.35
Japan	0.09	0.11	0.26	0.30
Year 20				
U.S.	1.56	1.84	0.87	1.00
Great Britain	1.43	1.35	0.82	0.77
Germany	1.23	1.35	0.62	0.67
Japan	0.87	0.99	0.45	0.50
Citations to Japan Citation Intensity Relative to Japan-Japan				
Year 1				
U.S.	0.16	0.20	0.27	0.34
Great Britain	0.15	0.17	0.26	0.29
France	0.15	0.17	0.25	0.31
Germany	0.18	0.21	0.30	0.38
Year 20				
U.S.	2.12	2.75	1.61	2.01
Great Britain	1.70	2.05	1.27	1.49
France	1.43	1.76	1.03	1.23
Germany	1.62	2.00	1.16	1.39

Inclusion of the technological proximity parameter has an effect similar to the exclusion of self-cites, but much smaller. That is, except for citations to the U.S., Column 3 shows slightly less localization than Column 4 (and Column 1 slightly less than Column 2), whereas both Columns 1 and 2 show dramatically less than either 3 or 4. What this suggests is that citations within the same patent class have a slight tendency to geographic localization, but, not surprisingly, much less so than citations within the same organization. Finally, there does not appear to be much interaction between the self-cite and technological proximity effects. The parameter γ is not much different in column 2 from in column 4. What this means is that self-citations exhibit approximately the same tendency toward concentration in the same patent class as non-self-citations.

V. CONCLUSIONS

In our view, the results in this paper demonstrate that there is much to be learned about international knowledge diffusion from patents and their citations. Despite the fact that we focus on patents granted by the U.S. patent office, rich patterns of interaction are revealed, including interesting findings about the diffusion of citations within and between countries other than the U.S. Some widely-held notions about differences in the inventive processes across countries were confirmed, such as the reliance of the Japanese on a relatively recent technological base. Others are less obvious, such as the strong symmetry between citing and cited intensities, and the greater proximity of Japan to the U.S. relative to Europe.

Overall, the results confirm our earlier findings that there is significant geographic localization of knowledge flows. We can now, however, tell a more complete story about the localization process, distinguishing the issue of speed from the issue of total intensity, and describing how citations diffuse over time to more distant locations. In future work, we intend to extend this in two directions. First, we will continue to look at more and finer geographic distinctions, including other countries and regions within the U.S. We conjecture, for example, that the West Coast of the U.S. is "closer" in technology space to the Pacific Rim, while the East Coast is closer to Europe, for both geographic and cultural reasons.

The second research avenue we are pursuing is to relate the knowledge flows implied by the citation patterns to the commercial impact of invention as measured by productivity improvements. If citations are a proxy for the pathways by which the cumulative impact of new technology is brought to bear, then they ought to play in a measurable intermediating role between the R&D series of various countries and the international productivity series. Thus our estimated citation flows can be used in place of trade flows to construct weighted stocks of foreign R&D to search for international R&D spillovers as in Park (1995) and Coe and Helpman (1995).

An issue that remains for further study is the extent to which the results may be tainted by systematic biases in the patent approval process that generates citations. Our maintained hypothesis is that the citation process itself does not differ depending on the domicile of the inventor. One possible bias is introduced by the fact that we are examining citations within the U.S. patent system. If a given invention is covered by patents issued in more than one country, then the obligation to cite this invention can be discharged by a citation to any of the members of the patent "family" around the world that cover the same invention in different countries. Further, U.S. inventions are often patented in the U.S. but not in Japan, while Japanese inventions patented in the U.S. are usually also patented in Japan. As a result, localization of citations to U.S. patents might be explained by a tendency of Japanese inventors to cite the Japanese patent covering prior art rather than the U.S. patent on the same invention, combined with the fact that such a patent will often be unavailable for U.S.-invented patents. This would not, however, explain why U.S. patents issued to Japanese inventors are *more* likely to cite other U.S. patents issued to Japanese inventors than they are to cite U.S. patents issued to *German* inventors; if anything, the bias introduced by patent families would suggest that our estimates of localization for citations to countries other than the U.S. are understated.

Our basic goal in this paper was to explore the process by which citations to a given patent arrive over time, how this process is affected by characteristics of the cited patent, and how different potentially citing locations differ in the speed and extent to which they "pick up" existing knowledge, as evidenced by their acknowledgment of such existing knowledge through citation. Recognizing that many inventions are never patented, that knowledge can flow from one inventor to another without being acknowledged by a citation, and that many citations probably do not reflect knowledge flow, we nonetheless view this process as a useful window into the otherwise "black box" of the spread of scientific and technical knowledge. The value of this view could obviously be enhanced, however, by a deeper understanding of the relationship between patent citations and knowledge flows. This will require more qualitative and institutional examination of inventions, patents and citations. A fruitful avenue is to build on the work of Jaffe, Fogarty and Banks by using inventors' detailed knowledge of the technological relationship between inventions to "test" the links implied by citations.

Patent citations offer a rich repository of information about the locus of technological activity, and the relationships among activities in different places. Systematic use of these data requires, however, careful attention to the need to control for time and technology field effects that otherwise have an impact on simple comparisons across countries or other units of observation. Fortunately, the patent data are sufficiently numerous that detailed controls can, in fact, be implemented. Though any model obviously imposes structure on the data, one can allow for complex patterns of interactions among effects. Indeed, readers of this paper have no doubt already thought of additional interactions that we could have estimated with our data but did not. We hope that, as these data become more widely available, other researchers will pursue questions that we have not considered.

REFERENCES

Caballero, Ricardo J. and Adam B. Jaffe (1993), "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," in Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1993*, Cambridge. MA: MIT Press.

Coe, David and Elhanan Helpman (1995), "International R&D Spillovers," *European Economic Review.*

Eaton, Jonathan and Samuel Kortum (1996), "Trade in Ideas: Patenting & Productivity in the OECD," *Journal of International Economics*, 40(3-4):251-278.

Griliches, Zvi (1979), "Issues in Assessing the Contribution of R&D to Productivity Growth," *Bell Journal of Economics*, 10:92-116.

Griliches, Zvi (1990), "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature,* 92:630-653.

Grossman, Gene M. and Elhanan Helpman (1991), "Quality Ladders in the Theory of Growth," *Quarterly Journal of Economics*, 106:557-586.

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

Jaffe, Adam B., Michael S. Fogarty and Bruce R. Banks (forthcoming), "Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation," *Journal of Industrial Economics.*

Jaffe, Adam B., Rebecca Henderson and Manuel Trajtenberg (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics,* 108:577-598.

Jaffe, Adam B. and Manuel Trajtenberg (1996), "Flows of Knowledge from Universities and Federal Laboratories," *Proceedings of the National Academy of Sciences,* 93:12671-12677, November 1996.

Krugman, Paul (1991), *Geography and Trade*, Cambridge, MA: MIT Press.

Park, Walter G., "International R&D Spillovers and OECD Economic Growth," *Economic Inquiry*, 33:571-591, October 1995.

Podolny, Joel and Andrea Shepard (1997), "When Are Technological Spillovers Local?: Patent Citation Patterns in the Semiconductor Industry," mimeo.

Putnam, Jonathan (1997), "How Many Pennies for Your Quote? Estimating the Value of Patent Citations," mimeo.

Romer, Paul M. (1990), "Endogenous Technological Change," *Journal of Political Economy,* 98:S71-S102.

Teece, D.J. (1977), "Technology Transfer by Multinational Firms: The Resource Cost of Transferring Technological Know-How," *Economic Journal,* 87:242-261.

Trajtenberg, Manuel, Rebecca Henderson and Adam B. Jaffe (1997), "University Versus Corporate Patents: A Window on the Basicness of Invention" *Economics of Innovation and New Technology.*

APPENDIX A

DERIVATION OF EXPECTED CITATION FREQUENCY FOR A "CELL"

Let s index the patent classes represented by patents with the t, ℓ, g attributes, and S index the patent classes represented in the set of patents with the T,L attributes. Let $N_{t g \ell s}$ represent the number of cited patents in a given class s, N_{TLS} the number of citing patents in a given class S, and $C_{t\ell gsTLS}$ be the total number of citations *from* class S in year T and country L *to* class s in year t, country ℓ and field g. Starting from Eq. **1**, the expected value of the citation count for a given "cell" is:

$$E[C_{t\ell gsTLS}] = (N_{TLS})(N_{t\ell gs})[1+\gamma D(s,S)]\alpha_{t\ell gTL}\exp[-(\beta_1)_{t\ell gTL}(T-t)][1-\exp(-\beta_2(T-t))] \qquad \textbf{(A1)}$$

where D(s,S) is now unity for s=S and zero otherwise. We can now sum over all s and S to yield:

$$E[C_{t\ell gTL}] = \sum_s \sum_S N_{TLS}\ N_{t\ell gs})[1+\gamma D(s,S)]\alpha_{t\ell gTL}\exp[-(\beta_1)_{t\ell gTL}(T-t)][1-\exp(-\beta_2(T-t)$$

or

$$E[C_{t\ell gTL}] = N_{TL}N_{tg\ell}\ \alpha_{t\ell gTL}\exp[-(\beta_1)_{t\ell gTL}(T-t)][1-\exp(-\beta_2(T-t))]\sum_s \sum_S [1+\gamma D(s,S) f_{TLS} f_{t\ell gs}]$$

where $f_{TLS}=(N_{TLS}/N_{TL})$and analogously for $f_{\ell gs}$. The double summation over s and S can be replaced by a single sum over s, because the only non-zero entries are where D(s,S) is unity or s=S. Thus

$$\frac{E[c_{t\ell gTL}]}{(N_{TL})(N_{t\ell g})} = \alpha_{t\ell gTL}\exp[-(\beta_1)_{t\ell gTL}(T-t)][1-\exp(-\beta_2(T-t))][1+\gamma\sum_s[f_{TLs}f_{t\ell gs}]] \qquad \textbf{2}$$

APPENDIX B

COMPLETE REGRESSION RESULTS

Parameter	Estimate	Asymptotic Standard Error	Asymptotic t-statistic
Technology Match (gamma)	99.489	2.903	34.3*
BETA1	0.190	0.002	107.5*
BETA2 ($x10^6$)	0.289	0.022	13.3*
ALPHAS:			
U.S. citing U.S.	1.000	n.a.	n.a.
U.S. citing Great Britain	0.710	0.013	-22.6
U.S. citing France	0.600	0.013	-30.0
U.S. citing Germany	0.545	0.011	-42.7
U.S. citing Japan	0.720	0.011	-25.9
Great Britain citing U.S.	0.722	0.014	-19.9
Great Britain citing Great Britain	1.781	0.031	25.2
Great Britain citing France	0.717	0.026	-10.8
Great Britain citing Germany	0.729	0.023	-12.0
Great Britain citing Japan	0.623	0.020	-19.3
France citing U.S.	0.654	0.014	-24.3
France citing Great Britain	0.791	0.028	-7.6
France citing France	2.170	0.037	31.2
France citing Germany	0.744	0.024	-10.8
France citing Japan	0.671	0.022	-14.8
Germany citing U.S.	0.560	0.012	-38.1
Germany citing Great Britain	0.746	0.023	-11.3
Germany citing France	0.726	0.024	-11.3
Germany citing Germany	1.320	0.022	14.5
Germany citing Japan	0.813	0.018	-10.4
Japan citing U.S.	0.761	0.012	-20.4
Japan citing Great Britain	0.659	0.020	-16.9
Japan citing France	0.634	0.021	-17.1
Japan citing Germany	0.827	0.018	-9.4
Japan citing Japan	2.318	0.025	53.1
BETA1s:			
U.S. citing U.S.	1.000	n.a.	n.a.
U.S. citing Great Britain	0.973	0.011	-2.4
U.S. citing France	0.967	0.014	-2.4
U.S. citing Germany	0.969	0.012	-2.5
U.S. citing Japan	1.052	0.010	5.0
Great Britain citing U.S.	1.033	0.012	2.6
Great Britain citing Great Britain	1.239	0.013	18.1
Great Britain citing France	1.083	0.025	3.3
Great Britain citing German	1.151	0.023	6.7
Great Britain citing Japan	1.094	0.023	4.1
France citing U.S.	1.033	0.014	2.3
France citing Great Britain	1.133	0.024	5.4
France citing France	1.306	0.014	22.0
France citing Germany	1.146	0.023	6.3
France citing Japan	1.163	0.025	6.4

Germany citing U.S.	1.033	0.013	2.5
Germany citing Great Britain	1.122	0.021	5.8
Germany citing France	1.125	0.023	5.4
Germany citing Germany	1.240	0.013	18.5
Germany citing Japan	1.180	0.017	10.6
Japan citing U.S.	1.190	0.011	17.1
Japan citing Great Britain	1.178	0.022	8.2
Japan citing France	1.164	0.024	6.8
Japan citing Germany	1.269	0.017	15.6
Japan citing Japan	1.543	0.010	54.3
ALPHAs:			
Drugs & Medical	1.000	n.a.	n.a.
Chemical, excl. Drugs	1.529	0.022	23.9
Electronics, etc.	2.279	0.031	41.8
Mechanical	1.857	0.025	33.9
All Other	1.765	0.032	24.2
BETA1s:			
Drugs & Medical	1.000	n.a.	n.a.
Chemical, excl. Drugs	1.018	0.008	2.3
Electronics, etc.	1.144	0.008	17.2
Mechanical	1.069	0.008	8.7
All Other	0.993	0.009	-0.8
CITING TIME PERIOD			
1963-65	1.000	n.a.	n.a.
1966-70	2.523	0.180	8.5
1971-75	4.048	0.288	10.6
1976-80	4.257	0.307	10.6
1981-85	3.936	0.289	10.1
1986-90	3.877	0.291	9.9
1991-93	3.276	0.253	9.0
CITING YEAR			
1977	1.000	n.a.	n.a.
1978	1.100	0.015	6.5
1979	1.094	0.016	6.0
1980	1.136	0.016	8.5
1981	1.152	0.016	9.4
1982	1.166	0.017	9.8
1983	1.167	0.017	9.6
1984	1.176	0.018	9.8
1985	1.228	0.019	11.8
1986	1.253	0.020	12.4
1987	1.334	0.022	15.0
1988	1.349	0.024	14.8
1989	1.355	0.025	14.4
1990	1.292	0.025	11.9
1991	1.266	0.025	10.7
1992	1.289	0.026	11.0
1993	1.298	0.028	10.8
1994	1.361	0.030	12.1

NOTES:

50625 observations

$R^2 = .7648$

Standard error of the regression = 7.55×10^{-5}

t-statistics are calculated for H_o:parameter = 1, except as noted

* t-statistic is for H_o:parameter = 0

Figures and Tables